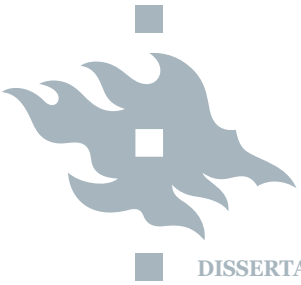

DISSERTATIONES SCHOLAE DOCTORALIS AD SANITATEM INVESTIGANDAM
UNIVERSITATIS HELSINKIENSIS

59/2016

SATU MÄKI-NEVALA

Molecular Markers in Finnish Lung Cancers

MEDICUM
DEPARTMENT OF PATHOLOGY
FACULTY OF MEDICINE
DOCTORAL PROGRAMME IN BIOMEDICINE
UNIVERSITY OF HELSINKI

Medicum
Department of Pathology
Faculty of Medicine
University of Helsinki
Helsinki, Finland

Molecular markers in Finnish lung cancers

Satu Mäki-Nevala

ACADEMIC DISSERTATION

To be presented with the permission of the Faculty of Medicine, University of Helsinki, for public examination in lecture hall 1, Meilahti Hospital, Haartmaninkatu 4, Helsinki, on October 14th, 2016, at 12 noon.

Helsinki 2016

Dissertationes Scholae Doctoralis Ad Sanitatem Investigandam Universitatis Helsinkiensis
No. 59/2016

ISSN 2342-3161 (print)
ISSN 2342-317X (online)
ISBN 978-951-51-2426-5 (paperback)
ISBN 978-951-51-2427-2 (PDF)

https://ethesis.helsinki.fi

Press: Hansaprint Oy, Turenki, 2016

Supervised by

Professor Sakari Knuutila, PhD
Medicum
Department of Pathology
University of Helsinki
Helsinki, Finland

Virinder Kaur Sarhadi, PhD
Medicum
Department of Pathology
University of Helsinki
Helsinki, Finland

Reviewed by

Vesa Kataja, MD, PhD
Chief Medical Director
Central Finland Health Care District
Jyväskylä, Finland
Adjunct Professor of Clinical Oncology
University of Eastern Finland
Kuopio, Finland

Docent Henna Tyynismaa, PhD
Research Program Unit - Molecular Neurology
University of Helsinki
Helsinki, Finland

Official Opponent

Docent Paavo Pääkkö, MD, PhD
Department of Pathology
Oulu University Hospital
Oulu, Finland

To the memory of my father

"Nothing in life is to be feared, it is only to be understood.
Now is the time to understand more, so that we may fear less."
— Marie Curie

CONTENTS

ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, which are referred to in the text by using their Roman numerals I–IV:

I. Mäki-Nevala S., Sarhadi V.K., Tuononen K., Lagström S., Ellonen P., Rönty M., Wirtanen A., Knuuttila A. and Knuutila S. Mutated ephrin receptor genes in non-small cell lung carcinoma and their occurrence with driver mutations – targeted resequencing study on formalin-fixed, paraffin-embedded tumor material of 81 patients. *Genes Chromosomes and Cancer*. 2013: 52:1141–9.

II. Mäki-Nevala S., Rönty M., Morel M., Gomez M., Dawson Z., Sarhadi V.K., Telaranta-Keerie A., Knuuttila A. and Knuutila S. Epidermal growth factor receptor mutations in 510 Finnish non-small-cell lung cancer patients. *J Thorac Oncol*. 2014: 9:886–91.

III. Mäki-Nevala S., Sarhadi V.K., Knuuttila A., Scheinin I., Ellonen P., Lagström S., Rönty M., Kettunen E., Husgafvel-Pursiainen K., Wolff H. and Knuutila S. Driver gene and novel mutations in asbestos-exposed lung adenocarcinoma and malignant mesothelioma detected by exome sequencing. *Lung*. 2016: 194:125–135.

IV. Mäki-Nevala S., Sarhadi V.K., Rönty M., Kettunen E., Husgafvel-Pursiainen K., Wolff H., Knuuttila A., and Knuutila S. Hot spot mutations in Finnish non-small cell lung cancers. *Lung Cancer*. 2016: 99:102–110.

ABBREVIATIONS

3'UTR	three prime untranslated region
5'UTR	five prime untranslated region
A	adenine
ADC	adenocarcinoma
ADSQ	adenosquamous cell carcinoma
ATP	adenosine triphosphate
bp	base pairs
C	cytosine
CI	confidence interval
CNB	core needle biopsy
CNV	copy number variation
COSMIC	the Catalogue of Somatic Mutations in Cancer
dbSNP	the Single Nucleotide Polymorphism Database
DNA	deoxyribonucleic acid
DNMTDNA	methyl transferase
EMT	epithelial-mesenchymal transition
Eph	ephrin receptor
FDA	the US Food and Drug Administration
FFPE	formalin-fixed, paraffin-embedded
FF	fresh frozen
FNA	fine needle aspiration
G	guanine
GTPase	small guanosine triphosphatase
HDAC	histone deacetylase
HR	hazard ratio
IARC	the International Agency for Research on Cancer
IHC	Immunohistochemical/immunohistochemistry
LCC	large cell carcinoma
LOH	loss of heterozygosity
MM	malignant mesothelioma
miRNA	micro-RNA
MPM	malignant pleural mesothelioma
mRNA	messenger-RNA
mtDNA	mitochondrial DNA
ncRNA	non-coding RNA
NGS	next generation sequencing
NSCLC	non-small cell lung cancer
NOS	not otherwise specified
OS	overall survival
PCR	polymerase chain reaction
PGM	Personal Genome Machine
PFS	progression free survival
RNA	ribonucleic acid

ROS	reactive oxygen species
RR	response rate
rRNA	ribosomal RNA
RTK	receptor tyrosine kinase
SCC	squamous cell carcinoma
SCLC	small cell lung cancer
SNP	single nucleotide polymorphism
T	thymine
SNV	single nucleotide variant
TKI	tyrosine kinase inhibitor
TSG	third generation sequencing
WHO	the World Health Organization

Gene names are used in according to guidelines of the Human Genome Organization Gene Nomenclature Committee (HGNC). Gene symbols are indicated in *italics* and those not listed here can be found at http://www.ncbi.nlm.nih.gov/.

ABSTRACT

Lung cancer is a common cancer with a poor prognosis. During the last decade, prognostic and predictive molecular markers for lung cancer have become available and treatment options have also multiplied. Targeted treatments have been developed aimed at distinct aberrant molecules driving the tumorigenesis. Thus, it is important to clarify the molecular characteristics of a tumor if one wishes to optimize the therapy. The currently available targeted therapies as well as further development in this field would be aided by the identification of (novel) significant markers and understanding how their incidence is associated with clinical characteristics. The aim of this thesis work was to examine known and novel molecular markers, more specifically mutations in selected genes which could be potential molecular markers of cancer, and to link these characteristics with clinical data in Finnish lung cancer patients. Thus, mutations were studied in genes encoding ephrin receptors, EGFR, and 22 other lung cancer related genes. In addition, protein coding genomic regions, i.e. exons, were studied to determine whether there were mutations associated with asbestos-exposure.

The study material consisted of more than 600 patients, their tumor specimens and clinical data. The majority of the specimens were formalin-fixed, paraffin-embedded non-small cell lung cancer (NSCLC) and malignant mesothelioma (MM) samples. Most of the specimens were subjected to next generation sequencing (NGS); the suitability of this technology in cancer diagnostics was also assessed. In particular, targeted and exome sequencing NGS methods were used with sequencing being performed on Illumina and Ion Torrent platforms. In addition, PCR-based mutation testing was used, and capillary sequencing being applied for validation purposes.

Mutations in ephrin receptor genes were common; 18 % of the patients carried one or more novel mutation. In MM, in particular EPHB1 was found to be mutated. The mutations did not associate with any particular clinical characteristic and they were found often concurrently with known pathogenic driver mutations, which points to a probable passenger mutation nature for these ephrin receptor mutations. However, when considering their diverse role in cellular function, as well as their oncogenic and tumor suppressive properties, therapeutically they may represent a very intriguing group of molecules. Thus, it would be important to clarify the significance of these alterations, especially at the mutation level.

Clinically significant EGFR mutations were found in 11 % of tumors from NSCLC patients. The mutations were associated with adenocarcinoma histology, female gender and never-smoking status, as has been reported in previous studies. The incidence of EGFR mutations resembled that described in previous studies conducted on other Western patients.

In the study comparing specimens from asbestos-exposed and non-exposed lung cancer, eight candidate genes (*BAP1*, *COPG1*, *INPP4A*, *MBD1*, *SDK1*, *SEMA5B*, *TTLL6* and *XAB2)* were found to be recurrently mutated exclusively in the asbestos-exposed patients. Candidate genes included those involved in cellular oxidative stress. Mutations in BAP1 and COPG1 were found exclusively in MM. BAP1 mutations and one SDK1 mutation were validated to be of somatic origin.

Screening of hot spot regions in 22 genes related to lung cancer revealed *TP53* and *KRAS* as the most frequently mutated genes, being mutated in 46 % and 26 % of the NSCLC patients, respectively. In particular, TP53 mutations were found to co-occur recurrently with

other mutations, also with pathogenic EGFR and KRAS mutations. Of the 425 patients, 77 % carried one or more mutations. Statistically significant associations were found between the following mutated genes and clinical characteristics: *TP53* and *PIK3CA* and squamous cell carcinoma, *KRAS* and adenocarcinoma, and *CTNNB1* and light ex-smoking status. The mutation profile was rather similar to that described in Western NSCLC patients with some exceptions, such as the higher BRAF mutation and lower STK11 mutation frequency.

The clinically significant mutations in Finnish NSCLC patients seem to resemble those detected in other Western patients. However, some differences can be found. Mutations in ephrin receptor genes are common and found often with other mutations. There seem to be molecular differences between asbestos-exposed and non-exposed lung cancers. However, the well-established lung cancer-related, pathogenic clinically relevant mutations, such as EGFR and KRAS, do not seem to be associated with asbestos-exposure. Finally, the application of NGS technology proved to be very suitable for cancer diagnostics. One major advantage of this technology is the possibility to test for different alterations in multiple genes simultaneously as well as the ability to detect and characterize both known and novel alterations.

TIIVISTELMÄ

Keuhkosyöpä on yleinen ja huonoennusteinen syöpä. Keuhkosyövän taudinkulun ennusteeseen (prognostinen) ja hoitovasteen ennusteeseen (prediktiivinen) vaikuttavista molekulaarisista markkereista tiedetään yhä enemmän, ja sen myötä syövän hoito on muuttunut. Käytössä on hoitomuotoja, jotka kohdentuvat tiettyyn syövän taustalla olevaan muuttuneeseen perimän molekyyliin. Kasvaimen molekulaaristen ominaisuuksien tunteminen mahdollistaa parhaan nykyaikaisen hoidon. Kohdennettujen hoitojen edelleen kehittäminen edellyttää tutkimusta merkittävien poikkeamien tunnistamiseksi, esiintyvyyden selvittämiseksi ja niiden yhteydestä taudin kliinis-patologisiin omaisuuksiin. Tämän väitöskirjatyön tavoitteena oli tutkia valikoitujen geenien uusia ja jo tunnettuja molekulaarisia markkereita, tarkemmin ottaen mutaatioita, suomalaisilla keuhkosyöpäpotilailla, ja niiden yhteyttä potilaiden kliinisiin ominaisuuksiin. Mutaatioita tutkittiin seuraavissa geeneissä: efriinireseptorit, *EGFR*, ja 22 keuhkosyöpään liittyvää geeniä. Lisäksi tutkittiin koko genomin proteiinia koodaavilta alueilta, eli eksoneista, asbestialtistukseen liittyviä mutaatioita.

Materiaalina tutkimuksessa oli yhteensä yli 600 potilaan aineisto sisältäen tuumorinäytteet ja potilaiden kliiniset tiedot. Valtaosa näytteistä oli formaliinilla fiksattuja, parafiiniin valettuja ei-pienisolukeuhkosyöpä (NSCLC)- ja mesotelioomanäytteitä (MM). Tutkimuksessa käytettiin pääsääntöisesti uuden polven sekvensointimenetelmiä (NGS), joiden luotettavuutta ja sopivuutta syöpädiagnostiikassa arvioitiin. NGS-menetelmistä käytettiin kohdennettua ja eksomisekvensointia. Sekvensoinnit tehtiin Illumina ja Ion Torrent-teknologioilla. Lisäksi käytettiin PCR-perusteista mutaatiotestausta, ja kapillaarisekvensointia tuloksien validoinnissa.

Mutaatiot efriinireseptori-geeneissä olivat yleisiä; 18 % potilaista todettiin vähintään yksi uusi mutaatio. Erityisesti EPHB1-mutaatiot toistuivat MM-potilailla. Mutaatiot eivät olleet kytköksissä mihinkään tiettyyn kliiniseen ominaisuuteen, ja ne esiintyivät usein yhdessä tunnettujen, patogeenisten aloitusmutaatioiden kanssa, mikä viittaa efriinireseptorimutaatioiden matkustajamutaatio-luonteeseen. Huomioiden efriinireseptorien monimuotoisen roolin solun toiminnassa, ja niiden sekä onkogeenisen ja tuumorisuppressiivisen potentiaalin, ovat ne hoidollisesti hyvin kiinnostava proteiiniryhmä. Siten niiden poikkeamien laajempi tunteminen myös mutaatiotasolla olisi tärkeää.

Kliinisesti merkittäviä *EGFR*-geenin mutaatioita löydettiin 11 % NSCLC-potilaista. Mutaatiot olivat kytköksissä adenokarsinooma-histologiaan, naissukupuoleen ja tupakoimattomuuteen, kuten aiemmissa tutkimuksissa on kuvattu. EGFR-mutaatioiden esiintyminen suomalaisilla potilailla muistutti vahvasti mutaatioprofiilia, joka niin ikään on aiemmin kuvattu länsimaalaisissa potilasaineistoissa.

Asbestialtistuneen ja altistumattoman keuhkosyövän vertailututkimuksessa löysimme kahdeksan kandidaattigeeniä (*BAP1*, *COPG1*, *INPP4A*, *MBD1*, *SDK1*, *SEMA5B*, *TTLL6* ja *XAB2)*, jotka olivat toistuvasti mutatoituneet vain asbestialtistuneilla potilailla. Kandidaattigeenit sisälsivät mm. solun hapetusstressiin liittyviä geenejä. BAP1 ja COPG1-mutaatiot löytyivät yksinomaan MM-näytteistä. BAP1-mutaatiot ja yksi SDK1-mutaatio validoitiin somaattisiksi.

Keuhkosyöpään liittyvien 22 geenin ns. hot spot-alueiden mutaatiokartoitus paljasti *TP53* (46 %) ja *KRAS*-geenit (26 %) aineistomme NSCLC-potilailla yleisimmin

mutatoituneiksi. Erityisesti TP53-mutaatiot esiintyivät toistuvasti useiden muiden mutaatioiden kanssa, myös patogeenisten, keuhkosyöpään liittyvien EGFR ja KRAS-mutaatioiden kanssa. Noin kolmella neljästä (77 %) NSCLC-potilaista (n=425) ilmeni vähintään yksi mutaatio. Tilastollisesti merkittäviä yhteyksiä löydettiin seuraavien geenimutaatioiden ja kliinis-patologisten ominaisuuksien välillä: *TP53* ja *PIK3CA* ja levyepiteelikarsinooma, *KRAS* ja adenokarsinooma, sekä *CTNNB1* ja kevyt tupakointihistoria. Mutaatioprofiili oli muutamin poikkeuksin hyvin samankaltainen kuin länsimaalaisilla potilailla aiemmin kuvattu. Tässä tutkimuksessa löydettiin korkeampi *BRAF*-mutaatio- ja matalampi *STK11*-mutaatiofrekvessi.

Suomalaisten NSCLC-potilaiden mutaatioprofiili muistuttaa vahvasti länsimaisissa potilasaineistoissa kuvattuja. Joitakin poikkeuksia kuitenkin löytyy. Asbestialtistuneen ja altistumattoman keuhkosyövän välillä vaikuttaa olevan molekulaarisia eroavaisuuksia. Hyvin tunnetut patogeeniset kliinisesti merkittävät mutaatiot, kuten EGFR and KRAS, eivät kuitenkaan näytä liittyvän asbestialtistukseen. Efriinireseptorien mutaatiot ovat yleisiä, ja ne esiintyvät usein muiden poikkeamien kanssa. Lisäksi, NGS-menetelmät sopivat hyvin syöpädiagnostiikkaan. Niiden ehdoton etu on mahdollisuus testata monenlaisia poikkeamia samankertaisesti, ja tunnistaa sekä jo tunnettuja että täysin uusia poikkeamia yksityiskohtaisesti.

1 INTRODUCTION

Cancer is one of the leading disease-burdens and a major cause of global mortality. It is characterized by rapid and uncontrolled cell growth. Cancer originates from an abnormal cell that has gained a growth advantage and managed to escape from the normal control system. Cancer may originate in any tissue, can grow in any part of the body, and spread to adjacent tissues or even to distant organs through metastases. One fundamental feature of cancer is that the malignant cells harbor an accumulation of genetic and epigenetic instabilities. These genetic alterations can be structural or numerical chromosome changes, smaller DNA sequence alterations, or epigenetic changes. However, only a minority of those alterations can actually drive tumorigenesis; these are considered as driver alterations, while other alterations are passenger events.

Conventionally cancer is treated with surgery, chemo- and radiotherapy. Nowadays, there are multiple targeted treatment options available for different cancer types. These treatments tend to be targeted against a certain aberrant molecule that is involved in cancer development and/or progression. However, unfortunately after a preliminary good response, in general, resistance develops to all of these treatments. The resistance may be attributable to several mechanisms, but it is believed that genetic alterations play a crucial role. Thus, it is important to characterize tumor molecular markers in detail, and to distinguish the significant driver alterations from the numerous passengers.

As our understanding of molecular basis of cancer has deepened, the tumor classifications have also supplemented details of location and cell morphology to include their molecular features. The revolution of genetic markers, their validation along with improvements of genetic diagnostics tools, has brought genomics-based cancer medicine into the clinics. The increasing numbers of targeted treatments demand that there are efficient tools for pinpointing those patients who will benefit from the targeted therapy. The first next generation sequencing (NGS) methods were introduced approximately a decade ago. After multiple and various development steps, improvements and validations, NGS has been approved for clinical use. They are particularly useful in diagnostics, as they enable to test in a time and cost-efficient manner multiple distinct alterations simultaneously from a small starting material.

Lung cancer is a fatal malignant disease with a very high incidence all around the world. Lung cancer may develop in any part of the lung. The most important cause of lung cancer is tobacco smoke. According to its histological features, lung cancer can be divided into two major groups: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). NSCLCs are the most common form, accounting for 85 % of all lung cancer cases. NSCLCs can be further divided into smaller subgroups based on cell morphology and molecular features. The characteristic molecular alterations encountered in adenocarcinoma tumors are especially used in the clinics, as some of these alterations may predict a response to targeted treatment.

Malignant mesothelioma (MM) is rarer, but a frequently fatal cancer associated strongly with asbestos exposure. MM originates in the mesothelium, the cell lining of internal organs. Most commonly MM develops in the linings of the lungs (pleura), abdomen (peritoneum) and heart (pericardium), in decreasing frequency. The molecular changes occurring in MMs are not as well-established as those in NSCLC, particularly the alterations associated with

asbestos-exposure (if any) remain obscure. The development of targeted therapies and predictive biomarkers would be one way to improve the outcome in MM.

It has become clear that ethnicity influences molecular changes. Thus, it was deemed relevant to study molecular markers in a national Finnish lung cancer cohort since the Finnish population has a history of genetic isolation. Therefore, this thesis focused on the investigation of molecular markers, i.e. mutations in DNA, in Finnish NSCLCs and MMs, especially their association with clinicopathological characteristics of the patients. The main methods used were NGS and PCR-based mutation testing.

2 REVIEW OF LITERATURE

2.1 Cancer genetics

Cancer is a complex malignant disease, expanding from an altered cell after its clonal expansion. The cell needs to possess certain genetic and epigenetic characteristics that confer on it an evolutionary benefit to undergo uncontrolled growth. Those genetic alterations may occur either in proto-oncogenes encoding the proteins involved in cell proliferation, differentiation, invasion and growth triggering an activation of those genes, or in tumor suppressor genes, which encode the proteins regulating cell cycle or DNA repair, causing their inactivation. In both cases, the alterations confer a growth advantage. In general, both alleles in the homologous chromosomes of tumor suppressor genes need to be altered, as they commonly act in a recessive manner. This "two-hit hypothesis" was first suggested by Knudson (1971) for retinoblastoma. On the contrary, proto-oncogenes are commonly dominant, thus an alteration in one allele can cause a gain-of-function change, and proto-oncogenes become transformed into oncogenes with pathogenic features.

In the review of Vogelstein et al. (2013) it was estimated that there may be from two to eight alterations present in a tumor that can actually drive the tumorigenesis (driver alterations), for instance, by altering pathways involved in cell fate, survival or genome maintenance. The others changes are considered as passenger alterations. Mutations in oncogenes are prone to occur in certain amino acid residues, whereas mutations in tumor suppressors occur in multiple distinct positions along the gene. It is known that lung cancer cells very frequently harbor somatic mutations; the median number of non-synonymous mutations being around 150 in a tumor (Vogelstein et al., 2013).

Genetic changes may occur at different levels, from small DNA sequence alterations to larger structural and numerical chromosome changes. If alterations occur in the protein-coding regions of the genome, i.e. exons, they may lead to the production of altered protein(s) encoded by a gene. Gene expression is regulated by several mechanisms that can also be altered. Numerous proteins are involved in the cellular signaling pathways that regulate many crucial cellular functions, such as programmed cell death, i.e. apoptosis, differentiation, proliferation and migration. It is these altered proteins and signaling pathways that serve as targets for novel therapeutic agents, such as monoclonal antibodies and small molecule inhibitors (reviewed in Ciavarella et al., 2010). The newest targeted agents are able to inhibit oncogenic features while preserving normal cellular function (Ciavarella et al., 2010).

Hanahan and Weinberg (2011) have described eight distinct hallmarks, known as the "Hallmarks of cancer", that a normal cell needs to acquire if it is to become malignant: 1) self-sufficiency in growth signals, 2) insensitivity to antigrowth signals, 3) evasion of apoptosis, 4) limitless replication potential, 5) sustained angiogenesis, 6) tissue invasion and metastasis, 7) reprogramming of energy metabolism, and 8) evasion of immune destruction. Genomic instability and inflammatory reactions of (pre)malignant lesions promote and contribute to the development of those hallmarks. Genomic instability provides a way that the tumor can gain new and probably more beneficial features to promote its growth. In particular, alterations in the genes involved in DNA repair enable an even higher accumulation of genomic instability. A cancer cell can harbor many different changes and

survive due to altered cell cycle regulators. (Hanahan and Weinberg, 2011) A tumor may be very heterogenic and consist of populations of genetically very different cells (de Bruin et al., 2014; Gerlinger et al., 2014). In addition to intra-tumor heterogeneity, metastases deriving from the primary tumor may be genetically different, and one metastasis can harbor different cell populations. Moreover, tumors between individuals are distinct. In other words, each cancer is unique displaying intra-tumor, inter-metastatic, intra-metastatic and inter-patient differences (Gerlinger et al., 2014).

Cancer predisposition is caused by both environmental and genetic factors. An individual may be exposed to different carcinogenic or mutagenic agents due to the environment in which he/she lives or the life-style, such as asbestos, air pollution, tobacco smoke, alcohol, unhealthy diet, too little physical activity, micro-organisms and exposure to sunlight. Furthermore, an individual's own genetic background plays a role, i.e this can either increase or decrease the risk for suffering cancer. (reviewed in Vineis and Wild, 2014) In conclusion, cancer development is a multifactorial and complex process, but altered genetics is always at its core (Hanahan and Weinberg, 2011).

2.1.1 Small DNA alterations

There are approximately 3 billion base pairs (bp) in the human genome. Although the sequence of the bases is a characteristic of humans, there are also often variations, such as single nucleotide polymorphisms (SNP). The DNA sequence can be altered once the cell undergoes cell division, i.e. mitosis, when the DNA structure is opened and undergoes replication. Small DNA changes include substitutions, small deletions and insertions, i.e. indels (Figure 1).

The smallest DNA alteration is the single nucleotide variation (SNV); this is a point mutation, where one nucleotide has become altered. In a substitution, one nucleotide base is substituted by another: in a transition, a purine nucleotide changes to another purine (A/G), and in a transversion, a purine changes to a pyrimidine (C/T). If mutations occur in an exon, they may alter the translation of the protein's amino acids. The mutation may be: 1) synonymous, silent mutation, if the amino acid remains the same, 2) nonsense mutation, if the mutation leads to a stop codon in the middle of the sequence, and protein translation is stopped prematurely causing a truncated protein product, or 3) missense mutation, if an amino acid is replaced by some other amino acid. Nonsense and missense mutations are also known as non-synonymous mutations, and they may lead to the synthesis of a defective protein.

DNA is transcribed in triplets of the nucleotides, i.e. codons, starting with a certain starting codon and ending with particular stop codons. Indels may cause a change in this reading frame and this can lead to an alteration called a frameshift mutation, or cause in-frame insertion/deletion of codons. Moreover, indels can alter splice sites, i.e. DNA sequence sites that are involved in correct removing of non-coding regions of the genes, i.e. introns.

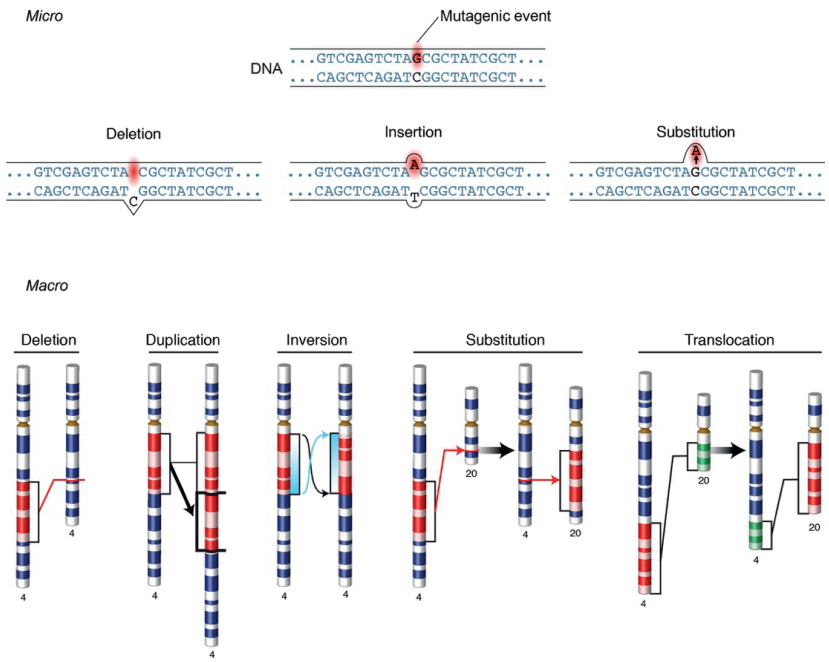

Figure 1. Genetic alterations. Micro-level DNA sequence alterations include small deletions and insertions, and substitutions. Structural chromosomal alterations are macro-level mutational events including deletions, duplications, inversion, substitution and translocations. Image courtesy: National Human Genome Research Institute, credit Darryl Leja, NHGRI.

2.1.2 Structural and numerical chromosome alterations

Genetic alterations may also include larger genomic changes. In many solid tumors, such as lung cancer, chromosomal changes are commonly present. These consist of deletions, duplications, inversions, substitutions, translocations (Figure 1) and changes in the chromosome number, i.e. aneuploidy (Vogelstein et al., 2013). The genome has also some particularly fragile sites that are prone to harbor translocations and deletions (reviewed in Durkin and Glover, 2007).

Deletions cause a loss of genetic material, they may even delete the whole gene(s). Duplications lead to copy number variations (CNVs), where the copy number of genes is elevated, and thus the encoded protein becomes overexpressed. In the normal cell, there are two copies of every gene, one allele in each homologous chromosome. Inversions, substitutions and translocations rearrange genetic material. Those may produce fused genes with oncogenic features or inactivate genes by truncating those or removing them so they are no longer regulated by their promoters. If a loss or a gain of genetic material occurs, the alteration is called unbalanced. If the whole chromosome is gained or lost, it is a numerical chromosome change.

2.1.3 Epigenetic alterations

Epigenetic changes are genetic alterations that do not alter the DNA sequence, but alter gene expression by other means. Examples of epigenetic changes are DNA methylation, histone protein modifications and expression of non-coding RNAs (nc-RNAs). The presence of epigenetic variability may account for the distinctly different risk for malignancies between individuals. Moreover, epigenetic alterations seem to be therapeutically important, and compounds inhibiting the activity of the enzymes regulating epigenetic events have been developed (reviewed in Morera et al., 2016).

DNA methylation regulates gene expression in the genome. It occurs at genomic regions enriched with cytosine and guanine bps, called CpG islands. DNA methyltransferases (DNMTs) are enzymes that add methyl groups covalently to the cytosine bases. In hypermethylation, the methyl groups are added to DNA causing the silencing of DNA sequence. Hypermethylation in promoter sequences suppresses the gene expression. Hypomethylation is the opposite event, where the methyl groups are removed from the cytosines, thus promoting the gene expression. DNA methylation is an important regulator of normal development since this requires a strict time and tissue-specific regulation of gene expression. DNA methylation is a mechanism involved in X chromosome inactivation that is needed for normal development of females, who have two X chromosomes (reviewed in Goldberg et al., 2007).

In cancer, methylation takes place commonly in the promoter regions of specific tumor suppressor genes, leading to their inactivation. Hypermethylation is commonly seen in the genes that harbor also somatic mutations, and are essential for the cell function, such as DNA repair, cell cycle control, motility and proliferation. However, hypomethylation also can be encountered in malignant tumors although it occurs less specifically and more likely at later stages of tumorigenesis. Hypomethylation may cause an activation of oncogenes and loss of imprinting. (reviewed in Langevin et al, 2015)

CpG islands are frequently mutated. For instance, the CpG island in the *TP53* gene harbor approximately 50 % of all somatic mutations in this gene (Rideout et al., 1990). Methylation may also increase carcinogenic effects, as has been shown in a case of some of the carcinogens present in tobacco smoke - acrolein (Feng et al., 2006) and benzo(a)pyrene diol epoxide (BPDE) (Yoon et al., 2001). If there is deamination, i.e. removal of an amine group, of methylcytosine, this can make this methylated base susceptible to a C>T transition point mutations, and transversion (G>T) mutations by promoting the effect of exogenous carcinogens (Langevin et al., 2015).

Histones are proteins involved in packing of genomic DNA into the chromatin structure. The basic structure of chromatin is called a nucleosome; this is formed by two of each histone types (H2A, H2B, H3 and H4), around which DNA is wrapped. Histones have N-terminal tails, which are available for post-transcriptional modifications. There are specific enzymes to catalyze these modifications e.g. methylation, acetylation, deamination, ubiquitylation, phosphorylation and sumoylation (Goldberg et al., 2014). These modifications alter the histones and, depending on the site of modifications, may further affect the chromatin structure and gene expression (Langevin et al., 2015). Histone modifications, and alterations in other proteins involved in those modifications, are clearly associated with the development of malignancy. For instance, deacetylation, i.e. removal of

an acetyl group from the histone tail leads to transcriptionally inactive DNA (Langevin et al., 2015).

Non-coding RNAs (ncRNAs) are a group of RNA sequences that are transcribed from the genes but are not translated into proteins. They can be divided into groups based on their size, from short to long ncRNAs. The most widely studied group is a type of short non-coding RNAs (sncRNAs) called microRNAs (miRNA) that are of ~18–22 nucleotides in length. They inhibit gene expression in a sequence-specific manner by binding to the complementary messenger-RNA (mRNA) molecule transcribed from the target gene (reviewed in Langevin et al., 2015, and Tuna et al., 2016). Dysregulation of miRNAs leads to altered expression of their target genes for example, in cancer, they may promote upregulation of oncogenes and inhibit tumor suppressors. The deregulation of miRNAs can be induced by mutations, deletions, copy number alterations, amplifications and epigenetic alterations (Tuna et al., 2016). Long non-coding RNAs (lncRNAs) (> 200 nucleotides) are also able to regulate gene expression, and possibly they are also involved in protein regulation and structural organization (reviewed in Shi et al., 2013, and Schmitt and Chang, 2016).

From therapeutic point of view, miRNAs are intriguing targets, and promising results have been obtained when they have been evaluated in cancer models. Inhibition of specific miRNAs has been found to repress oncogenic miRNAs, but to date, none of the therapeutic agents have been approved for clinical use. Further research will be required before this goal is achieved, especially to overcome delivery challenges and to decrease unwanted effects. (reviewed in Wen et al., 2015)

2.2 Lung cancer

2.2.1 Epidemiology

Lung cancer is the most incident type of cancer and a major cause of global cancer-related deaths, being responsible for up to 1.8 million newly diagnosed cases and 1.6 million deaths annually (Ferlay et al., 2015). This value represents 13 % of all new cancer cases and 19 % of all cancer-related deaths. Every year in Finland, there are approximately 2 500 new cases of lung cancer are diagnosed, two out of every three in male patients, and approximately 2 100 lung cancer-related deaths (Engholm et al., 2015). This means that lung cancer causes 11 % and 24 % in males, and 6 % and 12 % in females of all cancer cases and cancer-related deaths, respectively. In Finland, prostate cancer in male and breast cancer in female remain the most common cancer types (Engholm et al., 2015). The vast majority of lung cancers are diagnosed in patients older than 55 years of age (reviewed in de Groot and Munden, 2012).

Although cancer mortality rates have decreased in all those countries with reliable data, lung cancer remains very lethal (Hashim et al., 2016). The mortality rates are rising, particularly among female patients (Hashim et al., 2016) reflecting the increase in tobacco smoking incidence in women (de Groot and Munden, 2012). Tobacco smoke is the major cause of lung cancer, and specific subgroups of small cell carcinoma and squamous cell carcinoma are linked to cigarette consumption. In contrast, adenocarcinoma is the most

common subgroup in never-smoker patients, especially in women (reviewed in de Groot and Munden, 2012, and IARC, 2012a). It has been estimated that up to 85–90 % of all lung cancer cases are caused by tobacco smoke; however, only approximately 15 % of all smokers develop lung cancer (reviewed in de Groot and Munden, 2012, and Pallis et al., 2013). Tobacco smoke includes more than 70 known carcinogens that lead to the formation of DNA adducts and mutations (IARC, 2012a). Cessation of smoking lowers the risk for lung cancer (de Groot and Munden, 2012; IARC, 2012a), although the risk does not decrease to the level it was before the individual started to smoke (reviewed in Karam-Hage et al, 2014). Smoking not only elevates risk for lung cancer, but its component, nicotine, might also alter the resistance of the tumors towards radio- and/or chemotherapy (reviewed in Warren et al., 2013). Additionally, asbestos fibers, radon, micro-organisms, air pollution, such as polycyclic hydrocarbon (PAH) compounds, wood dust (Hancock et al., 2015) and genetic susceptibility have been linked to an increased risk for developing lung cancer (de Groot and Munden, 2012; Pallis et al., 2013).

2.2.2 Diagnosis and histopathology of non-small cell lung cancer

Lung cancer is diagnosed commonly only during the late stages of the disease. The diagnosis is based on clinical symptoms and lung imaging by radiography and/or computed tomography (CT) (reviewed in Ettinger et al., 2010, and Lemjabbar-Alaoui et al., 2015). A more detailed pathological and molecular diagnosis is made from resected sample or small biopsies obtained during surgical or diagnostic procedures (Ettinger et al., 2010). Tumors are staged at diagnosis by according to the tumor, node and metastasis (Union for International Cancer Control (UICC) TNM) classification which defines the size of primary tumor (T), cancer cell spread into the adjacent lymph nodes (N) and distant metastases (M) (Sobin et al., 2009). According to the latest edition, the stages are defined as Ia, Ib, IIa, IIb, IIIa, IIIb and IV. Stages Ia–IIIb represent local and locally advanced cancer (spread to lymph nodes), whereas stage IV describes cancer metastatic to other organs (Sobin et al., 2009; Lemjabbar-Alaoui et al., 2015). Staging is important as a prognosticator of the patient's outcome and it represents the basis for evaluating and planning the treatment (reviewed in Lemjabbar-Alaoui et al., 2015, and Tsao et al., 2016).

Although all lung cancers share the same origin, it is important to subgroup them correctly, as subgroups represent clinically different diseases (Travis et al., 2015). Lung cancer can be divided into two major groups based on their histopathological features: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC is the most common group accounting for to 85 % of all cases; its origin is in the lung epithelium, whereas SCLC derives from the hormonal neuroendocrine cells (Lemjabbar-Alaoui et al., 2015). SCLCs are undifferentiated and very aggressive cancers, which commonly spread into lymphatic vessels and lymph nodes and also to the brain (Lemjabbar-Alaoui et al., 2015). This thesis focused on NSCLCs. NSCLCs are a heterogeneous group of lung tumors and they can be divided further into smaller subgroups based on their characteristic histopathology (Lemjabbar-Alaoui et al., 2015; Travis et al., 2015). The largest subgroups of the NSCLCs are adenocarcinoma (ADC), squamous cell carcinoma (SCC) and large cell carcinoma (LCC) (Ettinger et al., 2010). Tumors of distinct subgroups may emerge through

molecularly different tumorigenesis pathways (reviewed in Kadara et al., 2016), and derive from pulmonary different sites: ADC is found in the peripheral sites, whereas SCC in centrally located in the major bronchi. LCC tends to be poorly differentiated and can be found in any parts of the lung (Lemjabbar-Alaoui et al., 2015).

NSCLCs are very complex tumors and the subgrouping needs to be done in great detail (Lemjabbar-Alaoui et al., 2015; Kadara et al., 2016). In 2015, the new 4th edition of World Health Organization (WHO) classification was published for lung tumors, the previous version dating from 2004 (Travis et al., 2015). The major changes included a recommendation to use immunohistochemistry (IHC) in addition to morphological features also for resected tumor samples. IHC markers can be used to distinguish tumors if the morphology is unclear, as it may well be the case with small biopsies and cytological samples. There are five IHC markers that are approved for use in the classification of NSCLCs, i.e. thyroid transcription factor 1 (TTF-1) and napsin-A, both with a sensitivity of 80 %, for ADCs; and P40, which is the most sensitive and specific, followed by P63 and cytokeratin 5/6 (CK5/6) for SCCs (Travis et al., 2015). The second major change was to include genetic testing in the diagnosis as a way of selecting some form of targeted therapy. Moreover, it was recommended that the group of LCC should include only undifferentiated tumors (lack of any morphological and IHC differentiation), others should be in different subgroups. A new classification for ADC as defined by the International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification (IASLC/ATS/ERS) of 2011 was also largely included. The revised version is based on tumor cell invasiveness: ADC in situ (AIS), minimally invasive (MIA) and invasive; and growth pattern: solid, lepidic, acinar, papillary and micropapillary predominances (Travis et al., 2011). However, due to the timing of this study, the 3rd edition of WHO classification from 2004 has been used in this thesis (Travis et al., 2004).

The correct classification is very important from a prognostic and therapeutic point of view, as further molecular testing and therapeutic options are based on the classification (Travis et al., 2015). For instance, patients with SCC show no or only a poor response to pemetrexed (Scagliotti et al., 2011), and bevacizumab is highly toxic since there is a high risk that it will cause severe pulmonary bleeding in SCC patients (Johnson et al., 2004). Moreover, resected ADC with micropapillary or solid predominance display an increased response to adjuvant chemotherapy compared to acinar or papillary predominant tumors (Tsao et al., 2016). Some molecular changes reflecting the sensitivity for targeted therapies, such as EGFR mutations and ALK fusions, are mostly found in ADCs and thus it is recommended that tumors with an ADC classification, and those in which ADC cannot be excluded, should be molecularly tested (Travis et al., 2015; Tsao et al., 2016).

2.2.3 Genomic alterations in non-small cell lung cancer

Lung cancer cells harbor defects in their regulatory systems. Transformation into a malignant cell is thought to occur in a multistep and sequence-specific process driven by the accumulation of genetic and epigenetic changes (Lemjabbar-Alaoui et al., 2015; Kadara et al., 2016). Lung tumors vary extensively in their genetics, and thus there are no identical tumors. Moreover, the tumors are heterogenic and consist of various subclones of malignant

cells harboring somatic mutations and many other alterations (de Bruin et al., 2014). For these reasons, it is very difficult to identify the most significant changes and those with driver capabilities (Vogelstein et al., 2013).

In recent years, major efforts have been expended in conducting comprehensive genetic screening studies. Some specific genetic alterations occur more often in ADCs, and on the other hand, others in SCCs (Cancer Genome Atlas Research Network 2012 and 2014; reviewed in Devarakondra et al., 2015). In ADCs, RTK/RAS/RAF pathway activation (76 %), PI3K-mTOR pathway activation (25 %), p53 pathway alteration (63 %), alterations of cell cycle regulators (64 %), alterations of oxidative stress pathways (22 %), and mutations of various chromatin and RNA splicing factors (49 %) are common (proportions of all cases). In a whole exome sequencing study of ADCs, eighteen genes were found to harbor at a significant rate somatic alterations of non-synonymous mutations, rearrangements or CNVs (Cancer Genome Atlas Research Network 2014). These significantly mutated genes included (proto-)oncogenes (*EGFR*, *KRAS*, *BRAF*, *MET*, *PIK3CA*), tumor suppressors (*TP53*, *STK11*, *KEAP1*, *NF1*, *RB1*, *CDKN2A*), chromatin modifying (*SETD2*, *ARID1A*, *SMARCA4*), RNA splicing (*RBM10*, *U2AF1*), transcription factor (*MGA*), GTPase (*RIT1*) genes. The somatic mutations were detected with a frequency of 8.87 mutations per mega base of DNA. The significant somatic amplifications were detected in the following genes: *NKX2-1, TERT, MDM2, KRAS, EGFR, MET, CCNE1, CCND1, TERC*, *MECOM,* in chromosomal region of 8q24 near *MYC,* and a novel peak containing *CCND3,* and deletions in *CDKN2A*. Moreover, hypermethylation was clearly observed in genes involved in the WNT pathway (Cancer Genome Atlas Research Network 2014). In a RNA-sequencing study, fusions were found to be present with the *ROS1*, *RET*, *PRKCB*, *NTRK*, *MET* and *ALK* genes (Stransky et al., 2014).

Similarly, pathways of NFE2L2/KEAP1 (34 %), squamous differentiation genes (SOX2/p63/NOTCH1*)* (44 %), PI3K/AKT (47 %), and CDKN2A/RB1 (72 %) have been found to be significantly altered in SCCs (proportions of all cases) (Cancer Genome Atlas Research Network 2012). The tumor suppressor gene *CDKN2A* was silenced by distinct mechanisms of methylation, mutation, exon skipping and deletion in 72 % of SCCs. Significantly mutated genes were: *TP53*, *CDKN2A*, *PTEN*, *PIK3CA*, *KEAP1*, *MLL2*, *HLA-A*, *NFE2L2*, *NOTCH1* and *RB1*, when considering all genes, plus *FAM123B*, *HRAS*, *FBWX7*, *SMARCA4*, *NF1*, *SMAD4*, *EGFR*, *APC*, *TSC1*, *BRAF*, *TNFAIP3* and *CREBBP*, when considering only genes annotated in the Catalogue of Somatic Mutations in Cancer (COSMIC). Those included (proto-)oncogenes (*PIK3CA*, *EGFR*, *BRAF*, *HRAS*), other membrane receptor (*NOTCH1*), tumor suppressors (*TP53*, *CDKN2A, PTEN, KEAP1, RB1, APC*, *TSC1*), chromatin modifying (*MLL2, SMARCA*, *CREBBP*), pathway regulators (*FAM123B, NF1*, *SMAD4*), ubiquitination (*FBWX7),* transcription factor in oxidative stress (*NFE2L2),* and immune system related (*HLA-A, TNFAIP3*) genes. Tumors were also characterized by a chromosomal 3q gain. The observed mutation frequency was 8.1 mutations per mega base of DNA (Cancer Genome Atlas Research Network 2012). In another study, fusions were detected in *PRKCB*, *PRKCA*, *PKN1*, *FGR*, *FGFR1*, *FGFR2* and *FGFR3* in SCC (Stransky et al., 2014).

Nowadays, the information of molecular changes of ADCs is being applied in the clinic (Travis et al., 2015). EGFR mutations and ALK fusions are studied, and specific inhibitors are administered to treat the patients with these particular alterations (reviewed in

Thunnissen et al., 2014, and Patel et al., 2015). Moreover, numerous studies are ongoing to find predictive molecular markers, and to develop new therapeutic agents to treat molecularly different tumors. Table 1 lists some of most common genetic changes and novel therapeutics agents targeted against those alterations.

Table 1. Molecularly altered targets of (novel) inhibitory agents in NSCLC. Based on Patel et al. (2015), Tsao et al. (2015) and Abramson (2016).

Targetable gene alteration	Alteration prevalence (%)[a]	Functionality of molecule	FDA Approved targeted drugs[b]	Potential drugs in clinical trials[c]
ALK fusions	2–7	RTK	Crizotinib, ceritinib, alectinib	
EGFR mutations	10	RTK	Erlotinib, gefitinib, afatinib, osimertinib	
BRAF mutations	2–3	Serine/threonine kinase	-	Vemurafenib (2) Dabrafenib (2) Combination with MEK inhibitor
ERBB2/HER2 amplification	2 (ADC)	RTK	-	Afatinib (2) Dacomitinib (2)
ERBB2/HER2 mutations	1–4	RTK	-	Afatinib (2) Dacomitinib (2)
KRAS mutations	26	GTPase	-	MEK inhibitors mTOR inhibition
MEK1 mutations	1 (ADC)	Serine/threonine kinase	-	Selumetinib (2) Trametinib (2) Cobimetinib (1)
MET amplification	3 (ADC)	RTK	-	Crizotinib (2) Cabozantinib (2)
NTRK1 fusions	2 (ADC)	RTK	-	Enterectinib (2) LOXO-101 (2) Cabozantinib (2) DS-6051b (1)
PIK3CA mutations	3–4	Lipid kinase	-	LY3023414 (2) PQR 309 (1)
RET fusions	2 (ADC)	RTK	-	Cabozantinib (2) Alectinib (2) Apatinib (2) Vandetanib (2) Ponatinib (2) Lenvatinib (2)
ROS1 fusions	2 (ADC)	RTK	-	Crizotinib (2) Cabozantinib (2) Ceritinib (2)

RTK, receptor tyrosine kinase

[a] If proportion is calculated out of a certain NSCLC subgroup, the subgroup is indicated in the brackets

[b] Approved in treatment of NSCLC

[c] Number in brackets indicates the current phase of the trial: 1, phase I; 2, phase II; 3, phase III

2.2.3.1 EGFR mutations

Epidermal growth factor receptor (*EGFR*) gene is located on chromosome 7q12. It encodes a transmembrane protein, EGFR, which belongs to ERBB family of the receptor tyrosine kinases (RTKs) (GeneCards). Like the other RTKs, EGFR is a cell surface receptor; it is activated by ligand binding which triggers the dimerization of the receptor molecules inducing autophosphorylation of its tyrosine amino acids in the kinase domain, which leads to the activation of downstream signaling pathways, RAS/RAF/MEK and PI3K/AKT/mTOR, regulating the cell proliferation, differentiation and apoptosis (Sordella et al., 2004).

EGFR mutations occurring in the intracellular protein kinase domain (exons 18–21) can cause the constitutive activation of the receptor molecule and thus activation of the whole downstream signaling pathway even without ligand binding (Figure 2). These mutations alter the ATP-binding pocket of the receptor, which makes possible its activation without ATP binding. There are two common activating mutations, a single bp substitution causing a change in amino acid in the protein from leucine to arginine at codon 858 (Leu858Arg) in exon 21, and deletions in exon 19. Other activating mutations with known clinical significance are insertions in exon 20 and various missense mutations, such as Ser768Ile, Gly719Ala/Ser/Cys, Thr790Met and Leu861Gln. (reviewed in Sharma et al., 2007) EGFR mutations are reported to be largely mutually exclusive with other driver alterations, such as KRAS, BRAF mutations and ALK fusions (Dearden et al., 2013; Gainor et al., 2013; Tissot et al., 2016).

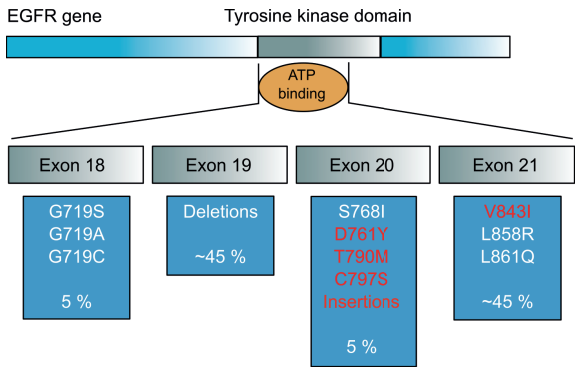

Figure 2. Activating EGFR mutations found in NSCLC. Mutations indicated in red predict insensitivity to EGFR-TKI therapy, whereas others (indicated in white) predict sensitivity. Adapted from Sharma et al. (2007).

EGFR mutations are found with a frequency of close to 10 % in European NSCLC patients (Gahr et al., 2013). In Asian NSCLC patients, the incidence is higher - as high as 50 % (Dearden et al., 2013). EGFR mutations are associated with ADC histology, female gender, never-smoking status and Asian ethnicity (Dearden et al., 2013). Certain activating EGFR mutations predict that the patient will respond favorably to small molecule receptor tyrosine kinase inhibitors (TKI), such as gefitinib, erlotinib and afatinib (Figure 2) (Mok et al., 2009; Zhou et al., 2011; Sequist et al., 2013). However, some of the mutations predict insensitivity to targeted drugs, and acquired mutations can cause resistance to targeted treatments after a good initial response (Sequist et al., 2011a and 2011b).

2.2.3.2 ALK fusions

Anaplastic lymphoma kinase (*ALK*) gene is located on chromosome 2p23 encoding the ALK receptor, which also belongs to the RTK family (GeneCards). Normally, the ALK receptor has an important role in the development of the brain, but alterations in this gene are found in cancer causing the downstream activation of RAS/MEK/ERK, STAT3 and PI3K/AKT pathways regulating cell proliferation, differentiation, survival and apoptosis (reviewed in Shaw and Solomon, 2011).

In lung cancer, ALK is activated by a translocation leading to constitutive dimerization of the receptor and induction of cell proliferation and survival via altered signaling pathways. Several fusion partners have been detected, such as *KIF5B*, *TFG* and *KLC1*, but the most common is a small inversion in the chromosome arm 2p, causing *ALK* to fuse with *EML4* (Figure 3). This *ALK-EML4* fusion occurs in exon 20 of *ALK* and it includes its tyrosine kinase domain, but *EML4* is truncated at distinct locations. (Soda et al., 2007; Takeuchi et al., 2009)

ALK fusions can be detected in approximately 3 % of unselected NSCLC patients. They are associated with ADC histology, never- or light smoking status and younger age, and are mutually exclusive with other driver mutations. The ALK-EML4 fusion predicts sensitivity towards ALK inhibitors, such as crizotinib and ceritinib. (Shaw et al., 2009; Sequist et al., 2011a; Zhao et al., 2015)

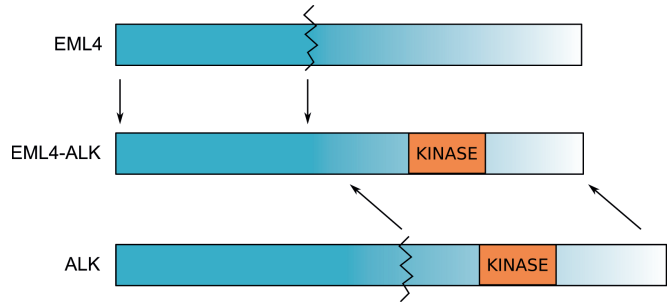

Figure 3. ALK-EML4 fusion in chromosome two arm 2p caused by a small inversion. Adapted from Soda et al. (2007).

2.2.3.3 Ephrin receptors

Erythropoietin-producing human hepatocellular receptor-interacting protein (ephrin) receptors (Ephs) form the largest group of RTKs. At present, gene family encoding the receptors has 14 members: *EPHA1-8*, *EPHA10*, *EPHB1-4* and *EPHB6*. Their ligands are ephrins, five in class A and three ephrin-B ligands. Some of the receptors have more than one ligand. Ephs differ in their signaling from other RTKs. Both a receptor molecule and a ligand are located on the cell surface, which leads to a contact-dependent cell interaction and bidirectional signaling. Forward signaling is dependent on Eph kinase activity, and reverse signaling on Src kinases. Moreover, the same cell can express both receptors and ligands, and even inhibit the signal induced by Ephs. In addition, some Ephs can act independently from ephrins. The crosstalk between different members of the large Ephs family and with other signaling pathway mediators, means that Ephs are crucial actors in cell signaling. The ephrin-mediated signaling is involved in various developmental events and cell homeostasis by regulating cell morphology, adhesion, proliferation, migration, survival and differentiation. The signaling outcome is highly dependent on cells and tissues. (reviewed in Pasquale, 2010, and Lisabeth et al., 2013)

Aberrations in Ephs and ephrins have been found in many cancers, and they have displayed both tumor suppressive and tumorigenic potential (Lisabeth et al., 2013). Ephs can promote signaling via Rho and Ras family GTPases in various pathways (Lisabeth et al., 2013). Intriguingly, Ephs can utilize those GTPases to suppress cell proliferation, survival and migration, this differs from the situation with the other RTKs using the same downstream signaling mediators (Pasquale, 2010; Lisabeth et al., 2013). Eph signaling promotes epithelial phenotype, and suppresses cell adhesion, migration and growth (Pasquale, 2010). In lung cancer, mutations have been found particularly in EPHA3 and EPHA5 (Davies et al., 2005; Ding et al., 2008) as well as alterations in the expressions of multiple Ephs (reviewed in Barquilla and Pasquale, 2015). They might also possess a prognostic role, such as high expression of EPHA8 in ovarian cancer (Liu et al., 2016a), and EPHB2 (Husa et al., 2016) in breast cancer predict poor survival. In addition, EPHA2 expression has been associated with a poor prognosis in many cancers, except lung cancer, as revealed in a recent meta-analysis (Shen et al., 2014). Germline CNVs of *EPHA3* have been linked to susceptibility to hereditary prostate cancer in Finland (Laitinen et al., 2016).

Since they are such a diverse group of RTKs, Ephs are an intriguing group from the therapeutic point of view, and multiple agents inhibiting or promoting Ephs are under the investigation (Barquilla and Pasquale, 2015). For instance, an EPHA2 inhibitor has shown good results in NSCLC cell lines in a preclinical study (Amato et al., 2014). NSCLC cells with increased CNV in EPHA3, EPHA5 and EPHA8 exhibited sensitivity to dasatinib, a wide-range kinase inhibitor (Sos et al., 2009). Aberrations in Ephs may also alter therapy outcome. Increased expression of EPHA2 can lead to resistance to trastuzumab in breast cancer (Zhuang et al, 2010) and vemurafenib in melanoma (Miao et al., 2015); EPHB3 has been demonstrated to promote resistance to radiotherapy in NSCLC cells (Ståhl et al., 2013); and polymorphisms in EPHA5 and EPHA6 are thought to play a role in chemotherapy (taxanes) toxicity in solid tumors (reviewed in Frederiks et al., 2015). Moreover, Ephs have been shown to participate in cancer-related epithelial-mesenchymal transition (EMT) (Li et al., 2014a), which is a mechanism allowing tumors to resist targeted therapies (reviewed

Uramoto et al., 2010; Chung et al., 2011). In summary, this makes Ephs a very interesting group of molecules.

2.2.4 Epigenetic alterations in lung cancer

In addition to the somatic changes in DNA, epigenetic alterations are very frequently encountered in lung cancer. Lung tumors express similar hypermethylated tumor suppressor genes as observed in other solid tumors (Langevin et al., 2015). There are many hypermethylated genes e.g. tumor suppressors of *RASSF1A*, *APC* and *CDKN2A,* and DNA repair gene *MGMT*. Multiple repetitive sequences, such as short and long interspersed nuclear elements (SINE and LINE), long transposable repeat (LTR) elements, duplicates and subtelomeric sequences, are frequently hypomethylated in SCC, as is commonly the case in other human cancers (Rauch et al., 2008). The methylation pattern seems to change in the course of tumorigenesis (Langevin et al., 2015).

There are also histone modification-related epigenetic changes in lung cancer; the HDACs are overexpressed, and a SIN3A involved in suppression of HDAC is downregulated, leading to closed chromatin assembly and subsequently suppression of gene expression, particularly of tumor suppressors (Langevin et al., 2015). Moreover, hyperacetylation of H4K5 and H4K8, hypoacetylation of H4K12 and H4K16, and loss of trimethylation of H4K20 have been detected in lung cancer (van den Broeck et al., 2008; Langevin et al., 2015).

NcRNAs are also altered in lung cancer, for instance, *miR-21* and *miR-210* seem to be more overexpressed than in normal lung, as shown in multiple studies (Guan et al., 2012; Võsa et al., 2013). *MiR-21* inhibits gene expression of *PTEN*, *PDC4* and *TPM1,* promoting angiogenesis and tumor growth in hepatocellular carcinoma (Meng et al., 2007). There are several miRNAs which have been reported to be methylated and thus inactivated (Langevin et al., 2015). The hypomethylation of one miRNA, *let-7a-3,* leads to the dysregulation of more than 200 genes, among those (proto-)oncogenes *RAS*, *MYC* and *HMGA2*, and thus can promote tumorigenesis (Brueckner et al., 2007). With respect to the lncRNA genes, *MALAT1* and *HOTAIR* are overexpressed, the former is thought to inhibit genes regulating metastasis and cell motility, while the latter is involved in regulating the histone regulatory complexes; both of these NcRNAs have highly conserved RNA sequences in mammals (Langevin et al., 2015).

Epigenetic alterations may also be relevant for therapeutics. Two inhibitors against the enzymes regulating epigenetic events have been tested in lung cancer i.e. DNMT and HDAC inhibitors (Langevin et al., 2015). However, at present they are not approved for treatment of lung cancer. Four HDAC inhibitors have been approved by the Food and Drug Administration (FDA) to treat lymphomas or myelomas (Abramson et al, 2016).

2.2.5 Standard treatment of non-small cell lung cancer

Lung tumors are largely treated with the standard methods: surgery, radio- and chemotherapy. If a tumor is found in its early stages, surgery is the most curative treatment option. The resection of lung tumors are large operations and cannot be conducted on patients in a very weak condition. More advanced NSCLCs are treated with a combination of radio- and platinum-based chemotherapies. Surgery can also be combined with radio- and chemotherapy. Neo-adjuvant therapy is given before the surgical operation to decrease the size of the tumor, and adjuvant therapy after the operation to destroy left-over malignant cells. (Lemjabbar-Alaoui et al., 2015; Tsao et al., 2016)

Lung tumors are most commonly treated with the standard methods: surgery, radio- and chemotherapy, or their combinations. If a tumor is detected while in its early stages, surgery is still the most appropriate treatment option, despite the emergence of "radio-surgery" with focused high-dose stereotactic radiotherapy. Resection of lung tumors is a major surgical procedure which cannot be performed on patients with comorbidities, as is often the case. In these cases, radiotherapy may be a better option. More advanced NSCLCs are treated with a combination of radio- and platinum-based chemotherapies. Neo-adjuvant therapy is given before the surgical operation to decrease the size of tumor, and adjuvant therapy after the operation to destroy any left-over malignant cells. In neo-adjuvant and adjuvant situations mainly radiotherapy and chemotherapy modalities are used. (Lemjabbar-Alaoui et al., 2015; Tsao et al., 2016)

2.2.6 Targeted treatment of non-small cell lung cancer

Targeted treatments have been developed; these are aimed at certain aberrant proteins or molecules in the cell driving the tumorigenesis (Thunnissen et al., 2014). The target should be measurable by a predictive biomarker that should monitor the clinical outcome after the applied treatment (Patel et al., 2015). At present, ALK and EGFR targeted therapies are approved to treat NSCLC patients with ALK fusion and activating EGFR mutations, respectively (Travis et al., 2015; Abramson, 2016). Those therapies improve the outcome of the patients harboring the alterations, compared to standard chemotherapy (Patel et al., 2015). For patients whose tumors harbor activating EGFR mutations, EGFR-TKI therapy (erlotinib, gefitinib or afatinib) is recommended (Lemjabbar-Alaoui et al., 2015; Travis et al., 2015). Those small molecule inhibitors inhibit EGFR signaling and induce apoptosis in cancer cells. Recently, immunotherapy (nivolumab) has been approved by FDA as treatment of advanced SCCs (Travis et al., 2015; Abramson, 2016).

However, regardless of all the success in the investigation of biomarkers and targeted treatments, only a fraction of patients with NSCLC can benefit from those therapies. Thus, SCCs and a majority of ADCs are still treated with the standard methods (Lemjabbar-Alaoui et al., 2015). In the future, identification of novel targetable alterations and new drugs, as well as determining the optimal combinatorial treatments are needed.

2.2.6.1 EGFR inhibitors

EGFR inhibitors consist of monoclonal antibodies (mAb) and small molecule inhibitors. The antibodies downregulate EGFR and subsequently inhibit its downstream signaling by competitively preventing the binding of ligand to the binding site of the receptor (reviewed in Pirker, 2015). These drugs are beneficial for NSCLC patients with high levels of EGFR expression in their tumors. Phase III randomized clinical trials have demonstrated the efficacy of the first generation antibody, cetuximab, and the second generation drug, necitumumab (Pirker, 2015). A meta-analysis claimed that cetuximab plus platinum-based chemotherapy could increase overall survival (OS) in advanced NSCLC patients compared to chemotherapy alone i.e. 10.3 months vs. 9.4 months (HR=0.88; 95 % CI, 0.79–0.97; p=0.009) (Pujol et al., 2014). Similarly, patients receiving necitumumab plus gemcitabine and cisplatin showed longer a OS of 11.5 months compared to 9.9 months for gemcitabine and cisplatin (HR=0.84; 95 % CI, 0.74–0.96; p=0.01) in advanced SCCs (Thatcher et al., 2015).

EGFR-TKIs are reversible or irreversible drugs, binding competitively to intracellular ATP-binding pocket of the tyrosine kinase domain of EGFR. TKI prevents ATP-binding and subsequently phosphorylation of EGFR, leading to downregulation of survival and proliferation signaling pathways (Sharma et al., 2007). In 2003, gefitinib was approved by FDA for previously treated advanced NSCLC (Cohen et al., 2003), and a year later, 2004, erlotinib received a similar approval (Cohen et al., 2005). In 2004, the first studies describing the association between the EGFR mutations and the efficacy of EGFR-TKIs were published: EGFR mutation positive NSCLC patients showed a clear sensitivity towards two first-generation EGFR-TKIs i.e. gefitinib (Lynch et al., 2004; Paez et al., 2004; Pao et al., 2004) and erlotinib (Pao et al., 2004). At present, both gefitinib and erlotinib are approved as first-line monotherapy for advanced (or metastatic) EGFR-mutation positive NSCLCs (Lemjabbar-Alaoui et al., 2015; Abramson, 2016). Moreover, second and even third generation EGFR-TKIs are available and approved as targeted therapy for advanced NSCLCs. Whereas the first generation drugs bind reversibly, the second generation drug, afatinib, forms irreversible covalent bonds with EGFR. Afatinib was developed to target a resistance-causing gatekeeper mutation, Thr790Met, but it did not overcome the resistance (reviewed in Minguet et al., 2016). The third-generation drugs, of which osimertinib is approved by FDA, are similarly irreversible and target also the gatekeeper mutation in addition to the activating EGFR mutations, while not binding to wild-type EGFR, which decreases their toxicity (reviewed in Tan et al., 2016).

In phase III randomized controlled trials, Mitsudomi et al. (2010) reported longer median progression free survival (PFS) for gefitinib-treated patients compared to cisplatin plus docetaxel treated: 9.2 months vs. 6.3 months (HR=0.49; 95 % CI, 0.34–0.71; p>0.0001). Similarly, PFS of 13.1 months for erlotinib compared to 4.6 months for gemcitabine plus carboplatin (HR=0.16; 95% CI, 0.10–0.26; p>0.0001) (Zhou et al., 2011), and 11.1 months for afatinib compared to 6.9 months for cisplatin plus pemetrexed (HR=0.58; 95 % CI, 0.43–0.78; p=0.001) have been reported (Sequist et al., 2013). Moreover, afatinib showed benefit in PFS over erlotinib as a second-line treatment for advanced SCC: 2.6 months vs. 1.9 months (HR=0.81; 95 % CI, 0.69–0.96; p=0.0103) (Soria et al., 2015). Furthermore, an

increase in OS over erlotinib was reported: 7.9 months vs. 6.8 months (HR=0.81; 95 % CI, 0.69–0.95; p=0.0077) (Soria et al., 2015).

Osimertinib has shown good results in phase I/II clinical trials and is now undergoing a phase III trial. It is also under investigation for combination with other (novel) therapeutic agents (reviewed in Pirker et al., 2016, and Wang et al., 2016). In a phase I trial, the median PFS for osimertinib treated T790M positive advanced NSCLC patients was 9.6 months (95 % CI, 8.3 – not reached) vs. 2.8 months (95 % CI, 2.1–4.3) in T790M negative patients (Jänne et al., 2015). Other third generation drugs, such as rociletinib, are also under further investigation, having shown promising results in advanced NSCLC with T790M mutations (Minguet et al., 2016).

Taken together, a recent meta-analysis which assessed 30 randomized controlled trials, showed EGFR mutation positive patients treated with EGFR-TKIs enjoyed a clear PFS benefit over chemotherapy or placebo treatment, both in first-line and second/third-line settings (Liu et al., 2016b). Lee's et al. (2015) meta-analysis produced similar results, achieving the greatest benefit in female, never-smokers and exon 19 deletions EGFR positive patients. Moreover, patients with exon 19 deletions were observed to have increased OS when treated with first-line irreversible EGFR-TKIs compared with those patients with exon 21 Leu858Arg (Kuan et al., 2015). It has been suggested that exon 19 deletion positive patients form their own entity. Yang et al. (2015) analyzed OS data of EGFR-mutation positive ADCs comparing afatinib and cisplatin-based chemotherapy, and observed longer OS longer in patients with exon 19 deletion treated with afatinib compared to pemetrexed-cisplatin: 33.3 months vs. 21.1 months (HR=0.54; 95 % CI, 0.36–0.79; p=0.0015) and similarly compared to gemcitabine-cisplatin: 31.4 months vs. 18.4 months (HR=0.64; 95 % CI, 0.44–0.94; p=0.023).

2.2.6.2 ALK Inhibitors

ALK inhibitors are also small molecule inhibitors that have displayed in PFS level efficacy in advanced ALK fusion positive NSCLC (Shaw and Engelman, 2013). At present, there are three FDA-approved ALK inhibitors available for clinical use: crizotinib, ceritinib and alectinib (Abramson, 2016), although only crizotinib is available in Finland. Ceritinib and alectinib have been claimed to function against brain metastases and gatekeeper mutations causing crizotinib-resistance, and are thus approved for treating metastatic, crizotinib-resistant, ALK positive disease (Khozin et al., 2015). In 2011, crizotinib was approved for the second-line treatment of advanced ALK positive NSCLCs, and two years later as a first-line treatment (Malik et al., 2014).

In a randomized phase III trial, crizotinib was found to be superior over pemetrexed/docetaxel chemotherapy in ALK fusion positive advanced or metastatic NSCLC: PFS 7.7 months vs. 3.0 months (HR=0.49; 95 % CI, 0.37–0.64; p<0.001) (Shaw et al., 2013). Similarly, the superiority of crizotinib over chemotherapy was reported in the first-line treatment for ALK positive non-squamous NSCLCs (Solomon et al., 2014). In a phase I trial, ceritinib showed good responses in advanced NSCLCs with ALK fusion either with or without crizotinib-resistance mutations, and thus it may be able to overcome crizotinib resistance (Shaw et al., 2014). Alectinib was recently approved after displaying

good responses and tolerability in a phase II trial in ALK fusion positive crizotinib-resistant NSCLC patients (Shaw et al., 2016). Further clinical trials with ceritinib and alectinib are ongoing (Minguet et al., 2016).

2.2.6.3 Angiogenesis inhibitors

Inhibition of angiogenesis is a recognized anti-cancer treatment (Minguet et al. 2016). However, in lung cancer, the treatment effect has been rather disappointing. In randomized phase III trials, bevacizumab, an anti-VEGF-A, prolonged PFS in advanced non-squamous NSCLCs when used in combination with chemotherapy (cisplatin plus gemcitabine) up to only 6.7 months compared to 6.1 months with only chemotherapy, i.e. around 2.5 weeks (HR=0.75; 95 % 0.62–0.91; p=0.003) (Reck et al., 2009). Similar results in PFS were reported a year later (Reck et al., 2010). Although the difference may be statistically significant, there is no true clinical value. In addition, bevacizumab poses an increased risk of potentially fatal hemorrhage, making the drug not well-suitable for treatment of lung cancer (Johnson et al., 2004). Other inhibitors against angiogenesis signaling are under investigation, one of the most promising agents is nintedanib. It has been claimed to be beneficial for ADC patients when used in combination with chemotherapy in comparison with chemotherapy alone (Reck et al., 2014).

2.2.6.4 Immunotherapy

Immunotherapy exploits the patient's own immune system to attack cancer cells by targeting signaling pathways that are used by malignant cells to escape from immune recognition. Targeting immune checkpoint inhibitors in PD-1/PD-L1 (programmed death-1 receptor/programmed death-1 ligand) pathway, or CTLA-4 (cytotoxic T-lymphocyte antigen 4) enables the immune system to recognize cancer cells and attack them (reviewed in Keir et al., 2008, and Farkona et al., 2016). Cancer cells express PD-L1 protein on their surfaces serving as a ligand for PD-1 receptor expressed on the surface of activated T-cells. The ligand-receptor interaction inactivates T-cells and inhibits the immune response (Keir et al., 2008). If one can block this interaction, T-cells may be activated and mount an immune attack against tumor cells (Keir et al., 2008). The immune checkpoint inhibitor, CTLA-4, is expressed on the surface of T-cells and after binding of its ligands, CD80 and CD86, it inhibits T-cell activity. By blocking CTLA-4, the immune response can be activated against tumor cells (Farkona et al., 2016).

In 2015, nivolumab, a PDL-antibody, was approved by FDA for the treatment of SCC NSCLCs, in progression after chemotherapy (Kazandjian et al., 2016). Approval was given after a phase III trial, in which nivolumab showed an increased response rate (RR), OS and PFS compared to docetaxel in advanced SCCs regardless of PDL-expression status (Brahmer et al., 2015). RR was 20 % for nivolumab vs. 9 % for docetaxel (p=0.008), OS was 9.2 months vs. 6.0 months (HR=0.59; 95 % CI, 0.44–0.79; p<0.001), and PFS 3.5 months vs. 2.8 months (HR=0.62; 95 % CI, 0.47–0.81; p<0.001), respectively. Pembrolizumab and MPDL3280A are two other promising agents targeting the same PD-1/PD-L1 pathway; they are currently undergoing clinical trials (Minguet et al., 2016). The

CTLA-4 checkpoint inhibitor, ipilimumab, has shown PFS benefit in advanced NSCLC over chemotherapy and placebo (p=0.05) in a phase II trial (Lynch et al., 2012), and a phase III investigation is ongoing.

2.2.6.5 Other targets

As only a fraction of all NSCLC patients may benefit from EGFR and ALK inhibitors, there is a need for drugs that could benefit these kinds of individuals (Patel et al., 2015). There are multiple possible targetable aberrant and/or deregulated molecules in NSCLC cancers, such as RTKs of HER2, MET, NTRK1, RET and ROS1, and downstream signaling molecules, such as AKT1, BRAF, MEK1, KRAS and PI3K. Some of the potential drugs and their investigations are summarized in Table 1 (based on Tsao et al., 2015; Patel et al., 2015, and Abramson, 2016).

2.2.6.6 Resistance to targeted EGFR and ALK treatments

Unfortunately, the patients treated with targeted EGFR and ALK inhibitors eventually exhibit disease progression and resistance develops to the inhibitors even if the patients have enjoyed a good preliminary response. Resistance occurs approximately about 12 months after the start of the treatment (Sequist et al., 2011b; Katyama et al., 2012). Although resistance mechanisms remain partly elusive, some of the mechanisms are well-known and others are plausible but not confirmed. Tumor heterogeneity creates challenges also in targeted treatment, as other alterations can be responsible for the resistance, and treatment can cause the clonal selection meaning that the cells capable of being resistant to treatment start to predominate (de Bruin et al., 2014; Bria et al., 2015). Resistance mechanisms can derive from the same or different genes, and activate the same or alternative signaling pathways, which can be caused by mutations, amplifications and/or over-expressions, various pathway activations, epigenetic changes, or even morphological changes (reviewed in van der Wekken et al., 2016).

There are some well known resistance mechanisms for EGFR-TKIs as is the case with other EGFR mutations (Figure 2), e.g. Thr790Met (Pao et al., 2005; Sequist et al., 2011b), Asp761Tyr (Balak et al., 2006) and Cys797Ser (Thress et al., 2015), PTEN deletions (Sos et al., 2009), mutations in PIK3CA (Sequist et al., 2011b; Tan et al., 2016) and KRAS (Tan et al., 2016), HER2 or MET amplification (Engelman et al., 2007; Sequist et al., 2011b; Yu et al., 2013), epigenetic changes (Sharma et al., 2010; Langevin et al., 2015), and morphological change from NSCLC to SCLC or EMT (Sequist et al., 2011b; Yu et al., 2013; van der Wekken et al., 2016). The gatekeeper mutation, Thr790Met, is found in approximately 50–70 % of the cases treated with the first-generation EGFR-TKIs (Sequist et al., 2011b; Arcila et al., 2011; van der Wekken et al., 2016). This mutation alters the ATP-binding pocket of EGFR and suppresses TKI binding (Yun et al., 2008). The mutation Val843Ile is responsible for resistance to afatinib (second-generation TKI), and Cys797Ser to third-generation TKI (Thress et al., 2015; Tan et al., 2016). Activation of another pathway can overcome inhibition of EGFR and lead to activation of downstream signaling. For instance, the MET amplification which triggers the activation of the ERBB3/PI3K pathway,

can be detected in approximately 20 % of the first-generation EGFR-TKI treatment-resistant cases (Bean et al., 2007; Engelman et al., 2007).

Similar mechanisms underlie ALK-TKI resistance. The best established mechanisms involve various gatekeeper ALK mutations, such as Leu1196Met and Gly1269Ala, ALK and KIT amplification (Katayama et al., 2012), ALK CNVs, mutations in KRAS, and activation of ERBB signaling, such as EGFR (Doebele et al., 2012; van der Wekken et al., 2016), and a very recently reported change from NSCLC to SCLC (Cha et al., 2015; Caumont et al., 2016; Fujita et al., 2016). The gatekeeper mutations are found in approximately 40 % of crizotinib-resistant NSCLC patients (Doebele et al., 2012). Second-generation TKIs, ceritinib and alectinib, have been developed which target also the gatekeeper (Muller et al., 2016), however, resistance to these compounds also emerges. Recent studies have revealed novel secondary mutations in ALK (Katayama et al., 2014; Ou et al., 2015) and overexpression/increased activation of HER3, NRG1 (HER3 ligand), IGFR-1R (Dong et al., 2016; Isozaki et al., 2016), ALK, MET and EGFR (Isozaki et al., 2016), as possible mechanisms of resistance.

As discussed above, novel targeted inhibitors can overcome some of the resistance mechanisms. Multiple options of combinational therapies are under investigation to prevent resistance (Patel et al., 2015). For instance, EGFR therapy could be combined with inhibitors against MET and VEGF (reviewed in Carrera et al., 2014), and it has been proposed that ALK positive patients could be treated with heat shock protein 90 (HSP90) inhibitors (Sequist et al., 2010; Socinski et al., 2013), since HSP90 is an essential chaperone protein for tumor growth and it stabilizes many mutant proteins including the ALK fusion product (reviewed in Pillai et al., 2014).

2.3 Mesothelioma

2.3.1 Epidemiology

Mesothelioma originates in the mesothelium, a thin layer of tissue surrounding body cavities. Mesothelioma can be found in the lining of the lung (pleura), abdomen (peritoneum) and heart (pericardium). The most common type of mesothelioma derives from the pleura accounting for approximately for 70 % of all mesothelioma cases, and the second common is peritoneal mesothelioma (~30 % of all cases). Malignant pleural mesothelioma (MPM) is a rare, but very aggressive type of cancer with a very poor prognosis. (reviewed in Yang et al., 2008, and de Assis et al., 2014a) This thesis focuses on MPM. In Finland, approximately 90 new cases are diagnosed annually, the vast majority (80 %) in males, and it is likely due to occupational exposure (Yang et al., 2008). The relative one year survival rate is approximately 50 %, while the 5 year survival rate is only 6 % (Engholm et al., 2015).

An exposure to asbestos is the most important cause for mesotheliomas; the first report of an association was described in 1960 by Wagner et al. in South Africa. It has been estimated that 80 % of mesotheliomas are linked with asbestos exposure (Yang et al., 2008). Asbestos is a group of naturally occurring hydrated mineral silicate fibers with extremely durable characteristics. Due to their features, their usage became common during the industrial era and they were widely exploited in manufacturing and construction until the 1970's (reviewed in IARC 2012b). The two major groups of asbestos are chrysotile (white asbestos) and amphibole (blue asbestos), the latter being highly linked to mesotheliomas, although the role of chrysotile has remained debatable (Yang et al., 2008; IARC 2012b). Inhaled asbestos fibers remain in the lung and evoke an inflammatory reaction e.g. involving TNF-α and NF-κB, leading to damage to DNA, and resistance to apoptosis. Asbestos can also cause a release of reactive oxygen and nitrogen species (ROS and RNS), which promote DNA damage. (Yang et al., 2008; Sekido, 2013; de Assis et al., 2014a)

Although the use of asbestos is now banned in many developed countries, its widespread use in the past century still is reflected in the incidence of MPM due to its very long latency time, as long as 50 years can elapse after the exposure until the development of the clinical disease. In addition to asbestos, genetic susceptibility, radiation, Simian virus 40 (SV40) and some other mineral fibers, such as erionite, have been associated with mesotheliomas. (Yang et al., 2008)

2.3.2 Diagnosis and histopathology

Although the latency time may be very long, the time from the onset to the diagnosis is short, as symptoms arise shortly after disease development. Due to the very aggressive nature of mesothelioma, it is often found at an advanced stage. Mesothelioma is also diagnosed by computed tomography (CT) and biopsies. Tumors may be classified further by using the TNM classification system. (reviewed in van Zandwijk et al., 2013)

Mesotheliomas are classified into three different groups based on their histomorphological features: epithelioid, biphasic and sarcomatoid. The classification remained the same in the new WHO Classification as it was in the 2004 edition (Galateau-

Salle et al., 2016). IHC markers are used to confirm the pathological diagnosis. It is particularly important to distinguish epithelioid mesothelioma from carcinomas. Calretinin, Wilm's tumor gene product (WT1) and podoplanin are considered as the most specific markers for mesotheliomas. It is recommended to use at least one cytokeratin, two mesothelial and two carcinoma-related markers to confirm the diagnosis. (van Zandwijk et al., 2013)

The prognosis differs among subgroups, as the mean OS rates are 18, 11 and 8 months for epithelioid, biphasic and sarcomatoid, respectively. In general, three determinants i.e tumor histology, cancer spread and the patient's performance status, are the main prognostic factors. (van Zandwijk et al., 2013; de Assis et al., 2014a)

2.3.3 Genetic alterations

Our knowledge of genetic alterations in MM has increased in recent years, but nonetheless, the possible asbestos-exposure associated molecular alterations have remained rather elusive (Kettunen and Knuutila, 2014). In the early cytogenetic study conducted by Tiainen et al. (1989), deletions and translocations in chromosome 1, and losses in chromosomes 1 and 4 were associated with high asbestos-exposure in mesothelioma. In mesotheliomas, frequent deletions in chromosome arms 1p, 3p, 4p, 4q, 6q, 9p, 13q, 15q and 22q can be found (Sekido, 2013). Those regions include tumor suppressor *CDKN2A/ARF* (9p21) and *NF2* (22q12). In addition to alterations in *CDKN2A/ARF* and *NF2*, also *BAP1* mutations are well-established in MM (Sekido, 2013; Ugurluer et al., 2016).

Cyclin-dependent kinase inhibitor 2A (CDKN2A)/alternative reading frame (ARF) encodes both a cyclin-dependent kinase inhibitor ($p16^{INK4}$) and an alternative protein called $p14^{ARF}$, both of which are involved in regulating the cell cycle (GeneCards). The former phosphorylates and inactivates retinoblastoma protein and thus arrests the cell cycle. The latter suppresses the cell cycle by inhibiting the degradation of p53. Alterations, especially deletions, of *CDKN2A* may lead to activation of the cell cycle (Sekido, 2013). *CDKN2A* is deleted up to 80–90 % of MMs (de Assis et al., 2014a).

Neurofibromatosis type 2 (NF2) encodes Merlin protein that is also involved in the regulation of the cell cycle, and its loss/inactivation leads to cell cycle progression (GeneCards). Merlin also represses Rac/Pak and FAK (focal adhesion kinase) signaling, which regulates cell migration (Poulikakos et al., 2006). Inactivation of *NF2* by deletions or mutations is found in approximately 40 % of MMs (Sekido 2013; de Assis et al., 2014a). NF2 may be inactivated also by other means, such as epigenetic regulation and phosphorylation (Sekido, 2013).

BRCA1-associated protein 1 (BAP1) is a tumor suppressor gene located in chromosome 3p21.3. The BAP1 protein interacts with many other proteins, and is involved in multiple cellular functions, such as histone deubiquitination, cell cycle, chromatin modification, gene transcription and DNA repair (GeneCards). Genetic susceptibility is caused by BAP1 germline alteration, which predisposes to a cancer syndrome (Carbone et al., 2012). In addition, somatic BAP1 mutations occur in MMs with an approximated frequency of 20 % (Bott et al., 2011; Zauderer et al., 2013).

As discussed, RTKs are frequently altered (activated) in cancer and similarly in MM. For instance, EGFR and AXL, and also SRC family kinases have been found to be aberrantly activated (Sekido 2013). In contrast, *TP53* is one of the most frequently mutated genes in cancer, but it seems to harbor mutations much less frequently in mesothelioma (Sekido 2013; de Assis et al., 2014b). Recent sequencing studies have revealed a genome-wide allelic loss as well as frequent *SETDB1* mutations (Kang et al., 2016), mutations and CNVs in *CUL1* (Guo et al., 2015), and mutations in *SMACB1*, *PDGFRA*, *KIT*, *KDR*, *HRAS*, *PIK3CA* and *STK11* genes (Lo Iacono et al., 2015).

2.3.4 Treatment

MM is highly resistant to chemo- and radiotherapy and thus its prognosis remains poor (van Zandwijk et al., 2013). Cisplatin plus antifolate chemotherapy is the standard treatment for mesothelioma (Vogelzang et al., 2003; van Meerbeeck et al., 2005), but there are no guidelines for second-line treatment once disease progression occurs. Since the disease is normally only detected at a late stage, combined with its tendency to grow along the cavity/organ surfaces without forming distinct tumors, this means that surgery is rarely an option (van Zandwijk et al., 2013). Surgery has been found to significantly improve patient survival in individuals with malignant peritoneal mesothelioma (Miura et al., 2014).

Multiple targeted agents have been studied, but results have been unimpressive, and only a few drugs have been tested in phase III trials. Of those, VEGF (bevacizumab and thalidomide) and HDAC inhibitors (vorinostat) have been investigated in clinical phase III trials (reviewed in Hiddinga et al., 2015). Bevacizumab in combination with chemotherapy (pemetrexed plus cisplatin) increased OS of MPM patients compared to chemotherapy alone: 18.8 months vs. 16.1 months (HR=0.77; 95 % CI, 0.62–0.95; p=0.0167). However, bevacizumab evoked also adverse effects (Zalcman et al., 2015). Thalidomide as second-line treatment for patients whose disease had progressed on chemotherapy, did not improve PFS compared to active supportive care: 3.6 months vs. 3.5 months (HR=0.95; 95 % CI, 0.73–1.20; p=0.72) (Buikhuisen et al., 2013). Similarly, vorinostat as second or third-line treatment did not achieve any improvement in OS in comparison to placebo: 30.7 weeks vs. 27.1 weeks (HR=0.98; 95 % CI, 0.83–1.17; p=0.86) (Krug et al., 2015).

There are some novel targeted drugs under investigation i.e. podoplanin antibodies, CARP-1 inhibitor and cilengitide (reviewed in Hiddango et al., 2015). Podoplanin, a transmembrane sialomucin-like glycoprotein, is highly expressed in MPM, and its antibody has been found to efficiently inhibit cell growth in MPM cells (Abe et al., 2013). CARP-1 is a perinuclear phosphoprotein, which regulates cell proliferation and apoptosis signaling, and its inhibition by functional mimetics was reported to induce anti-metastatic events in MPM cells (Jamal et al., 2014). Cilengtide inhibits cell growth in MPM cells (Cheng et al., 2014). Understanding the molecular background and aberrations of mesothelioma in more detail is important for the development of better therapies and their combinations (Kettunen and Knuutila, 2014).

2.4 Methods in molecular cancer research

There have been huge advances made in the analytical methods for studying cancer-related genetic and epigenetic alterations in the last decade(s). Here, the most basic and widely-used methods will be presented briefly, with the major focus on NGS methods, as these were the main techniques exploited in this thesis. The selection of method depends on which alterations need to be assessed as well as the analytical sensitivity of the method and the material from which it is to be analyzed.

2.4.1 Conventional methods

Polymerase chain reaction (PCR) was developed in 1983 as a way to amplify DNA by utilizing a temperature-resistant DNA polymerase, synthetic DNA primers, free nucleotides and cycles of temperature changes (Mullis, 1990). Nowadays, numerous genetic applications rely on PCR. For example, in mutation testing, there are PCR-based commercial, customized kits available to amplify only the mutated sequence by using specific primers for mutations or blocking the wild-type sequence amplification. Other widely used methods are based on hybridization, i.e. the capability of DNA sequence to bind to a complementary sequence. Comparative genomic hybridization (CGH) has been exploited to investigate CNVs based on comparison of genomic DNA against the reference genome (Kallioniemi et al., 1992). In situ hybridization (ISH) makes possible also spatial localization of the sequence in the original tissue by utilizing specific labeled probes that bind only highly complementary sequences (Langer-Safer et al., 1982). Microarray techniques exploit multiplexing and use specifically arranged probes on an array (Chang, 1983). IHC is another important technique based on the use of labeled specific antibodies and the immunological properties of the molecules (Coons et al., 1941).

2.4.2 DNA sequencing

DNA sequencing methods started slowly evolving after the discovery of DNA structure by Watson and Crick in 1953. The technique requires the presence of a template DNA sequence that is produced according to the complement rules by using DNA polymerase and free (modified) nucleotides. The first method to become widely used, nowadays referred to as first generation sequencing, was Sanger sequencing (direct sequencing) published in 1977 (Sanger et al., 1977). In Sanger sequencing, some of the free nucleotides are dideoxynucleotides (ddNTPs) at a low concentration, ddNTPs terminate the DNA elongation (due to a lack of free 3’ hydroxyl group). One type of ddNTPs is included in the sequencing reaction at one time, producing various lengths of DNA fragments which can be separated by electrophoresis and grouped by ddNTPS. During the following decades, many improvements have occurred e.g. the sequencing became automated, and the automated capillary sequencer was developed (reviewed in Heather and Chain, 2016). Capillary sequencing was introduced as a shotgun method (Figure 4); this approach was used to sequence euchromatin of the human genome, of which 99 % was published in 2004 (International Human Genome Sequencing Consortium, 2004).

The pyrosequencing method represented an important development; this is a direct sequence-by-synthesis technique similar to the Sanger procedure. Instead of modified nucleotides, this approach was based on the luminescence produced by pyrophosphate (PPi) synthesis that occurs naturally once a nucleotide is incorporated. Two enzymes are used in pyrosequencing: a pyrophosphate is converted into ATP by ATP sulfurylase, and ATP is used as a substrate by luciferase, which produces a luminescent signal detected by camera which can be analyzed by a software program. The signal intensity is proportional to the numbers of nucleotides that are incorporated. (Ronaghi et al., 1998)

Subsequently, the molecular methods have developed extensively, for example, first generation sequencing methods have been largely replaced with the next generation or even the third generation sequencing methods (reviewed in Metzker, 2010, van Dijk et al., 2014, and Heather and Chain, 2016).

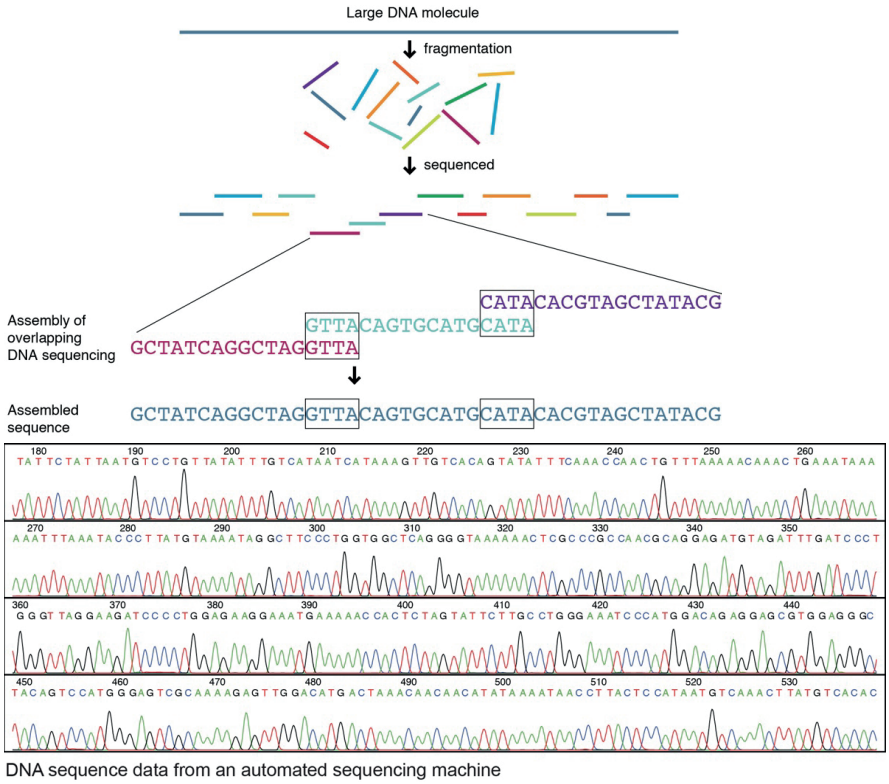

Figure 4. Shotgun sequencing is a method in which DNA is randomly fragmented, sequenced and partly overlapped sequences are organized in the correct order by computation. Automated capillary sequencing produces a chromatogram with fluorescent peaks of four colors, each for one type of nucleotide base (A, T, C or G). Image courtesy: National Human Genome Research Institute, credit Darryl Leja, NHGRI.

2.4.2.1 Next generation sequencing

The first NGS method, massively parallel signature sequencing, was published in 2000 (Brenner et al., 2000), followed by 454 Life Sciences sequencing commercialized by Roche, and the technology developed in 2004 by Margulies et al. (2005). One common feature of all methods is that they have a solid surface where the millions of sequencing reactions occur simultaneously. All methods are based on a template preparation from genomic DNA, sequencing, detection of results and alignment to the reference genome. NGS methods have multiple applications. In genomic DNA sequencing, whole genome, exome, targeted and *de novo* sequencing are used. NGS has revolutionized research investigating the connection between genetic variation and phenotype. Recently, a new application for NGS methods was developed - it can be utilized in a single cell genomics, for instance, making it possible to classify tumor cells according to their genetic features. This provides a deeper understanding of tumor development. In addition to genomic sequencing, RNA sequencing is available for investigating post-translational modifications, splice variants, gene fusions and changes in gene expression. This provides much larger and detailed information compared to array technology. (Metzker, 2010; van Dijk et al., 2014; Heather and Chain, 2016) In this thesis, DNA sequencing methods of targeted and exome sequencing have been used on Illumina and Ion Torrent platforms; these will be discussed in more detail in the following section. An overview of the platforms is described in Figure 5.

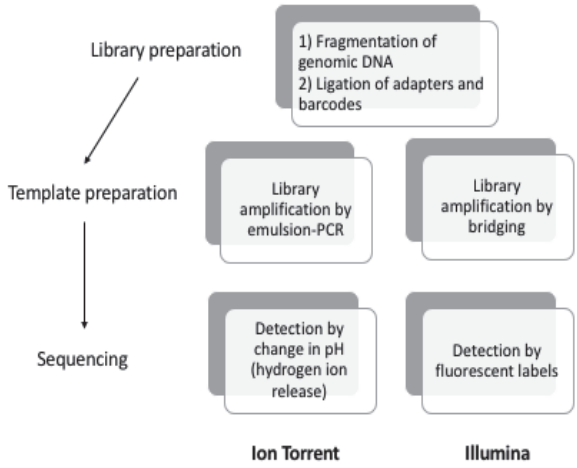

Figure 5. Overview of NGS workflow from the genomic DNA to sequencing on Ion Torrent and Illumina platforms. (By the author)

After the 454 Life Sciences (Roche), Solexa (Illumina) (Bentley et al., 2008), new techniques were published such as SOLiD (former Life Technologies) (Valouev et al., 2008) and Ion Torrent (Thermo Fisher Scientific) (Rothberg et al., 2011). The first step in all NGS methods, is the preparation of the library and the template, where the target genomic DNA is fragmented, and primers, adapter and barcode oligonucleotides are added to enable

specific sequencing chemistry and multiplexing. In amplicon-based NGS, such as the Ion Torrent technology, the libraries are size-selected and only those of the correct size are amplified by using primers designed for the genes of interest. The Illumina platforms (others than amplicon sequencing in Study III) used in this thesis, exploits an in-solution hybridization-based selection, where the specific biotinylated complementary probes select the library fragments to be amplified. The amplification may take place in emulsion (emulsion-PCR) or on a slide. The latter approach is used in Illumina and SOLiD. In the Illumina platform, it is performed by bridge amplification, where single stranded library fragments attach to the surface of the flow cell, which also includes primer sequences. In bridge amplification, the library molecules are amplified once they reach and bind primer sequences. The subsequent amplification event occurs close to the primary site, eventually forming a group of amplified templates. In emulsion-PCR, exploited by Ion Torrent, SOLiD and 454, amplification occurs in a mixture of emulsion oil and PCR mix. Ideally, one emulsion droplet contains one library molecule, primers, one bead and the PCR mix. After PCR cycles, a bead is covered by multiple copies of one type of library molecule. The library of nucleic acid fragments acts as a template for a new DNA to be synthetized. (Metzker, 2010; van Dijk et al., 2014; Heather and Chain, 2016)

NGS methods differ in the way they detect the nucleotides which have been incorporated into the elongating DNA sequence. The methods currently in use are reversible terminator sequencing (Illumina), semiconductor sequencing (Ion Torrent), pyrosequencing (454) and sequencing by ligation (SOLiD). Reversible terminator sequencing relies on the presence of fluorescently labeled nucleotides that terminate the sequencing reaction. All types of those nucleotides are applied in the same cycle and non-incorporated nucleotides are washed away. Fluorescent labels are detected by a camera in each flow cell. Before the next cycle, fluorescent dye and 3' terminator are removed. After all of the imaging cycles, the sequence is stored as fluorescent colors and mapped to the corresponding bases. (Metzker, 2010; van Dijk et al., 2014; Heather and Chain, 2016)

Semiconductor sequencing does not require imaging, but relies on the detection of a change in pH once a nucleotide has been incorporated and a hydrogen ion released. There is a sensitive sensor on the bottom of each micro well that contains the DNA template. A single nucleotide type is provided at a time, and if the nucleotide is complementary to the template, incorporation occurs and an ion release is detected. If more than one nucleotide is incorporated, a higher electronic signal is detected. (van Dijk et al., 2014; Heather and Chain, 2016)

The development of NGS methods poses enormous challenges for data management. Therefore, sophisticated bioinformatics tools for the data analysis are required. Figure 6 illustrates the basic workflow of NGS data analysis. NGS methods have proven to be more accurate and sensitive at detecting genetic alterations than the more conventional methods, such as FISH (Pekar-Zlotin et al., 2015) and Sanger sequencing (Chevrier et al., 2014; Han et al., 2014; Belardinilli et al., 2015). Due to high sensitivity and requirement of only a small amount of starting material, also archival material (FFPE) and fine needle aspirations can be tested with NGS (Hadd et al., 2013; Lin et al., 2014; Belardinilli et al., 2015).

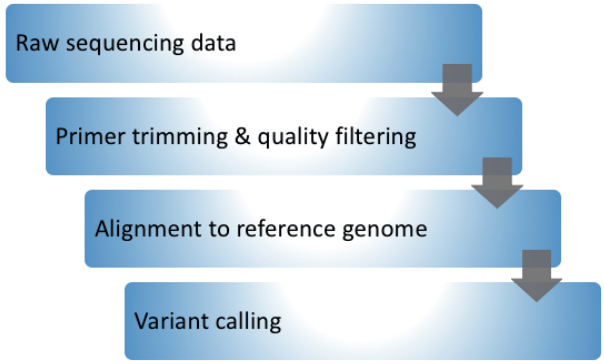

Figure 6. Overview of NGS data analysis steps. (By the author)

2.4.2.2 The third generation sequencing

Helicos (Helicos Genetic Analysis System), PacBio (Pacific Biosciences) and Oxford Nanopore Technologies (Oxford Nanopore Technologies Limited) represent the third generation sequencing (TGS) techniques that enable single molecule sequencing in real-time (SMRT) (Braslavsky et al., 2003; Harris et al., 2008; Pushkarev et al., 2009), whereas NGS requires the amplification of single molecules. TGS uses sequencing-by-synthesis technology and real-time detection of the fluorescent signal produced by the incorporated nucleotides. DNA polymerase initiates the sequencing reaction and the reaction ends when the polymerase dissociates from the template. The major advantage of TGS technology is the elimination of problems related to PCR and the short read length. Moreover, neither complex repeats nor poorly amplified regions pose a problem for the third generation technology. (reviewed in Roberts et al., 2013; Heather and Chain, 2016)

3 AIMS OF THE STUDY

The general aim was to study cancer-related mutations, their prevalence, concurrences and associations with clinical data, in Finnish NSCLC and MM patients, and to confirm the feasibility of utilizing next generation sequencing methods in cancer diagnostics.
The specific aims were to study:

- Mutations of ephrin receptor genes and their incidence in NSCLC (I)
- Clinically relevant epidermal growth factor receptor mutations in NSCLC(II)
- Novel asbestos-related mutations in lung ADC and MM (III)
- "Hot spot" mutations in *AKT1, ALK, BRAF, CTNNB1, DDR2, EGFR, ERBB2, ERBB4, FBXW7, FGFR1, FGFR2, FGFR3, KRAS, MAP2K1, MET, NOTCH1, NRAS, PIK3CA, PTEN, SMAD4, STK11,* and *TP53* in NSCLC (IV)

4 MATERIALS AND METHODS

The sample material and methods used in this thesis are only described generally in this section (Table 2). The original publications provide more detailed information.

4.1 Non-small cell lung cancer samples (I, II, III, IV)

FFPE tumor samples were collected from NSCLC patients diagnosed and treated in the Helsinki and Uusimaa Hospital District, Finland, during the years 2005 to 2014 (I, II and IV). In Study I, 24 samples were obtained from other hospitals in Finland. The majority of the patients had undergone surgical procedures. A minority of the samples were obtained by core needle biopsy (CNB) or fine needle aspiration (FNA). All the specimens were collected before any treatment, and had been evaluated by histology and tumor cell percentage by a pulmonary pathologist. Primarily all the samples tested had a tumor percentage of 20 % or more. The patient samples selected for Study I had an indication for ALK or EGFR testing. In Studies II and IV, samples were collected unselectively: all the samples available with sufficient tissue material, extracted DNA and tumor cell content were included into this study. In Study III, NSCLC samples represented fresh frozen (FF) tumor material. Normal tissue samples were used for validation purposes and were obtained from blood, normal adjacent lung or lymph nodes.

Diagnosis was based on the 3rd edition of WHO Criteria (Travis et al., 2004). In Study II, EGFR mutation positive ADC cases were furthermore sub-grouped according to the updated IASLC/ATS/ERS classification (2011) (Travis et al., 2011). Based on smoking history, patients were divided into the following subgroups: never-smokers, ex-light smokers (smoked <20 years but subsequently ceased >10 years ago), ex-medium smokers (smoked >20 years but subsequently ceased), and current smokers (smoked >20 years and ongoing). Evaluation of occupational asbestos exposure was based on the patient interview conducted in Studies I, II and partly in IV.

Table 2. Overview of the samples and methods used.

	Study I	Study II	Study III	Study III	Study IV
Tumor type	NSCLC	NSCLC	NSCLC	MM	NSCLC
Sample material	FFPE	FFPE	FF	FFPE	FFPE
Number of patients (original/results)	81/81	528/510	26/26	21/21	442/425
Female (%)	54	47	23	0	44
Never/Ever-smokers ratio[a]	9/53	67/438	1/25	9/12	38/383
Subgoups of NSCLC					
ADC (%)	91	78	100		67
SCC (%)	1	12	0		21
LCC (%)	5	8	0		8
Other (%)	2[b]	3	0		4
Methods used	Targeted NGS (Illumina); capillary sequencing; real-time PCR	Real-time PCR	Exome sequencing (Illumina); targeted amplicon sequencing (Illumina)	Exome sequencing (Illumina); targeted amplicon sequencing (Illumina)	Targeted NGS (Ion Torrent)

[a] Those in whom smoking history was known

[b] A diagnosis of one sample was changed to MM (epithelioid)

4.2 Malignant mesothelioma samples (III)

MM samples were also collected in the Helsinki and Uusimaa Hospital District, Finland, between 2007 and 2010, before any treatments. Similarly, the histological diagnosis and tumor cell content were investigated by a pulmonary pathologist. In total, 21 samples were eligible for inclusion in the study. Of those, all presented epithelioid histology.

4.3 DNA extraction (I, II, III, IV)

In all studies, DNA was extracted from FFPE tumor sections by using the QIAamp DNA Mini Kit (Qiagen GmbH, Hilden, Germany) with minor changes (Tuononen et al., 2013). The quality and quantity of the extracted DNA samples were measured with the Qubit® fluorometer (Thermo Fisher Scientific, Waltham, MA, USA).

In Study II, WaxFree™ DNA Extraction Kit (Trimgen, Sparks Glencoe, Maryland, USA) according to the manufacturer's instructions for paraffin samples was used for the isolation from some FNB and CNB FFPE specimens.

In Study III, DNA from FF specimens was extracted by using the QIAamp DNA Mini Kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer's protocol for extraction from FF tissue material.

4.4 Asbestos fiber measurement (III)

In Studies III and partly IV, the asbestos fibers were counted from normal lung tissue samples by scanning electron microscopy (SEM) from ADC specimens and by transmission electron microscopy (TEM) from MM samples. That was conducted according to the standardized method (Tuomi et al, 1992) in the Finnish Institute of Occupational Health, Helsinki, Finland. Criteria for patients to be considered as non-exposed were defined as the following asbestos fiber contents of dry lung tissue: less than $0.2x10^6$/g and $1.0x10^6$/g for MM and ADC, respectively. Asbestos-exposed lung contained fibers more than $1.0x10^6$/g and $2.0x10^6$/g in MM and ADC, respectively.

4.5 Real-time PCR (I, II)

In Study I, BRAF, EGFR and KRAS mutation testing was performed by standardized PCR-based mutation testing kits. EGFR mutations (19 distinct deletions in exon 19, three distinct insertions in exon 20, and point mutations of Gly719Ser/Ala/Cys, Ser768Ile, Leu858Arg and Leu861Gln) were tested using the TheraScreen EGFR PCR Kit (Qiagen®, Manchester, UK), KRAS mutations (Gly12Ala/Asp/Arg/Cys/Ser/Val and Gly13Asp) by the TheraScreen KRAS PCR Kit (Qiagen®), and BRAF V600E mutations by the AmoyDx™ BRAF V600E Mutation Detection Kit (Amoy Diagnostics, Xiamen, China). In all analyses, an ABI7500 instrument (Applied Biosystems, Foster City, CA, USA) was used.

In Study II, EGFR mutations were investigated by the TheraScreen EGFR PCR Kit (Qiagen®; 500 patients) or the cobas® EGFR Mutation Test (Roche Molecular Systems, South Branchburg, NJ, USA; 28 patients) on the cobas z480 platform. All the kits were used according to the manufacturer's instructions.

The Therascreen mutation tests use ARMS® (Newton et al., 1989) and Scorpions® technologies (Whitcombe et al., 1999; Thelwell et al., 2000). The AmoyDx kit is based on patented technology.

4.6 Sanger sequencing (I)

In Study I, capillary sequencing was performed on 13 samples with an Eph mutation and eligible material for mutation validation. DNA (5–15 ng) was amplified by PCR, and capillary sequencing was conducted both on tumor tissue and the corresponding normal tissue material of lymph node (eligible from 13 patients). Sequencing was performed on ABI 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA) using Big Dye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Warrington, UK). Results were analyzed on Sequencher 5.1 (Gene Codes Corporation, Ann Arbor, MI, USA).

4.7 Next generation sequencing

4.7.1 Targeted next generation sequencing on Illumina HiSeq2000 (I)

In Study I, 13 ephrin receptor genes, *EPHA1-8*, *EPHB1-4* and *EPHB6,* were analyzed for all exonic regions by targeted NGS. The gene panel had a total of 192 genes, including RTK, miRNA and cancer related genes. The target capture consisting of 2 676 regions (~1 Mb) was designed by using e-array (Agilent Technologies, Santa Clara, CA) to cover all exons, 3'UTR and 5'UTR regions of all the genes, and the RNA probes were obtained from Agilent.

Extracted DNA (2–3 μg) was prepared for the sequencing libraries by fragmentation, adapter ligation and target enrichment according to the SureSelect in-solution (Agilent) target capture and enrichment protocol based on hybridization and bridge amplification. The libraries were sequenced by paired-end sequencing on Illumina HiSeq2000 sequencer (Illumina, San Diego, CA, USA).

4.7.2 Exome sequencing on Illumina HiSeq2000 (III)

In Study III, DNA (1–3 μg) samples were prepared according to the NimbleGen SeqCap EZ Exome 2.0 Library SR User's Guide (Roche) with a few changes to target exonic regions. The Covaris S-2 instrument (Covaris, Woburn, MA, USA) was used for fragmentation. Both sample types, FFPE and FF were treated similarly. The libraries were sequenced on Illumina's HiSeq2000 sequencer.

4.7.3 Amplicon deep sequencing on Illumina MiSeq2000 (III)

In Study III, the amplicon deep sequencing was used for validation of the selected novel variations detected by the exome sequencing. PCR amplification was performed on DNA (10–20 ng) from paired tumor and normal adjacent lung tissue (if available), adapters were ligated, and PCR amplicons were sequenced as paired-end reads on Illumina's MiSeq2000 instrument.

4.7.4 Targeted next generation sequencing on Ion Torrent (IV)

In Study IV, DNA samples (10 ng) from 425 NSCLC samples underwent library preparation by using the Ion AmpliSeq™ Colon and Lung Cancer Panel (Thermo Fisher Scientific) together with the Ion AmpliSeq™ Library Preparation Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. The Colon and Lung Cancer panel covered 504 mutational hot spots (14.6 kb) in 22 genes: *AKT1, ALK, BRAF, CTNNB1, DDR2, EGFR, ERBB2, ERBB4, FBXW7, FGFR1, FGFR2, FGFR3, KRAS, MAP2K1, MET, NOTCH1, NRAS, PIK3CA, PTEN, SMAD4, STK11,* and *TP53*. The libraries were quantified with Qubit® 2.0 Fluorometer. Template preparation and enrichment were performed by using the Ion OneTouch™ 2 system (Thermo Fisher Scientific) and the Ion PGM™ Template

OT2 200 Kit (Thermo Fisher Scientific). The sequencing was conducted with the Ion PGM™ Sequencing 200 Kit v2, the Ion 316v2 chip and the Ion Torrent PGM instrument (Thermo Fisher Scientific).

4.7.5 Primary data analysis (I, III, IV)

In Studies I and III, the raw NGS data obtained by Illumina HiSeq sequencer were analyzed by using an in-house built analysis tool called variant calling pipeline (VCP), developed at the Finnish Institute of Molecular Medicine (FIMM) (Sulonen et al., 2011). The VCP is a combination of commonly used and in-house algorithms which can quality-filter the data, align the sequences utilizing the Burrows-Wheeler Aligner (BWA) (Li and Durbin et al., 2010) against the reference human genome (h19), remove duplicates, and call variants using SAMtools' (Li et al., 2009) for SNVs and Pindel (Ye et al., 2009) for indels. In Study III, prior to alignment, the overlapping paired reads were merged into single longer reads using SeqPrep. Only those variants with a read depth ≥6, a phred quality score ≥20 and a quality ratio of ≥0.8 were selected into the analysis. For indels, only those with a quality score ≥40 were included.

In Study III, the amplicon sequencing data were analyzed at FIMM by using an in-house algorithm pipeline including Bowtie2 (Langmead and Salzberg, 2012) to align the sequences against the reference (h19), call SNVs with SAMtools' and BCFtools, and indels with GATK IndelRealigner (DePristo et al., 2011).

In Study IV, the raw data obtained from the Ion Torrent PGM sequencer were analyzed with the TorrentSuite™ Software (v.4.0.2) (Thermo Fisher Scientific). The Variant Caller plug-in (v4.0-r76860) (Thermo Fisher Scientific) was used for variant calling with the default settings: a quality score ≥6, relative read quality ≥6.5, coverage ≥6 for SNP/COSMIC variant and 15 for indel, and strand bias ≤95 % for SNP/COSMIC variant and 90 % for indel. The Coverage Analysis plug-in (v4.0-r77897) (Thermo Fisher Scientific) was used for coverage analysis.

4.7.6 Secondary data analysis (I, III, IV)

The secondary analysis was performed by diverse steps. In the visualization of data, the Integrative Genomics Viewer (IGV) was used (Thorvaldsdottir et al., 2013). Two in silico analysis tools, PROVEAN (Choi et al., 2012) and SIFT (Kumar et al., 2009), were used to predict the effect of the non-synonymous variants on the encoded protein.

In Studies I, III and IV, all non-synonymous and indel variants in exonic regions were selected for further analysis. The variants reported in NCBI dbSNP (build 135 in Study I, build 137/142 in Study III, and build 146 in Study IV) were defined as known variants, as well as those not found in either dbSNP or 1000 Human Genomes Project as novel variants. In Study III, variants detected in any normal sample were also removed from the analysis.

In Study III, those mutations or genes that were detected exclusively in asbestos-exposed patient samples were selected out of all non-synonymous mutations causing a probable damaging effect. The exome data were investigated to find the most frequently mutated chromosomal positions and the genes involved. The recurrent variants or recurrently

mutated genes were defined as those occurring in three or more asbestos-exposed patients. All variants were checked in the dbSNP (build 142) to remove variants reported in a newly built database. Moreover, the exome data were studied also for variants (<2 % in the 1000 Genomes) found in *BRAF*, *EGFR*, *ERBB2*, *HRAS*, *KRAS*, *MET*, *NRAS, PIK3CA*, *STK11, EPHA1-8*, *EPHA10* and *EPHB1-4*, *EPHB6* for both ADC and MM, plus *BAP1*, *CUL1*, *CDKN2A*, *NF2* for MM. From amplicon deep sequencing data, the variants with a variant base frequency of 0.5 % of all reads in a given position were selected. The base frequency was compared to the quality value of the corresponding base. All variants with a frequency ratio of ≥0.7 were considered as true variants.

Study IV, a COSMIC mutation was defined as a mutation found in the Catalogue of Somatic Mutations in Cancer. In the secondary analysis, the following filtering threshold was set: a frequency of variant allele ≥3 %, and an average phred quality score ≥15 for previously known COSMIC mutations and ≥20 for novel variants.

4.8 Statistical tests (II, IV)

All statistical analyses were performed using IBM SPSS Statistics v21 (Study II) or v22 (Study IV) (IBM Corp., Armonk, NY, USA) software. The investigation of associations between clinicopathological features and molecular status was conducted using Pearson's χ^2 test, Fisher's exact test and independent t-test, as appropriate. In Study II, OS was estimated by the Kaplan-Meier method with a log rank test to compare the cumulative OS between the groups. Two-sided p values <0.05 were considered statistically significant.

4.9 Ethical issues

Ethical permissions for the studies were obtained from the Ethics Committee for the Department of Medicine of the Hospital District of Helsinki and Uusimaa, Finland; consent no. 370/13/03/01/2010.

5 RESULTS AND DISCUSSION

5.1 Ephrin receptor mutations in lung cancer (I, III)

In total, we examined 106 NSCLC and 22 MM samples for mutations in Ephs by NGS targeting the exons of the following Eph genes: *EPHA1-8*, *EPHB1-4*, *EPHB6* (Studies I and III), and *EPHA10* on 26 ADC and 21 MM samples in Study III. The stringent thresholds used in variant calling produced most likely true-positive variants, although some low-frequency variants may have remained undetected. In Study I, the average coverage was good with a mean of 135 (range 52–300). Moreover, our previous study showed that this NGS method produces reliable results compared with the golden standard mutation detection methods (Tuononen et al., 2013). In Study III, exome sequencing produced a mean target coverage of 38 (range 13–54), and similar thresholds were used.

When using the newest dbSNP (build 147) as the reference, the total number of novel non-synonymous mutations detected in both studies was 27 mutations in 23 patients. Of these, four cases were diagnosed with MM, 18 with lung ADC, and one with lung LCC. Of the 23 cases with mutations, 74 % (17/23) were males, and even more, 95 % (21/22; with known smoking status) were ever-smokers. In Study III, the mutations were detected in both asbestos-exposed (n=8) and non-exposed (n=5) patients (the asbestos-exposure status was known for 13 patients with a novel Eph mutation). The detected Eph mutations did not display any associations with the clinicopathological features. However, the sample set was small and thus firm conclusions cannot be drawn. Previously no association between mutations and clinical data has been conducted. The overall (novel) mutation frequency was 18 % (23/128) (Table 3), which is similar to that of the 20 % (37/188) found in a larger mutation screening experiment reported by Ding et al. (2008). The 21 rare variants (with the reported frequency less than 2 %), were found in 25 patients (including two MMs), accounting for 20 % (25/128) of the patients (Table 3). The Ephs were found to be mutated with similar frequencies in both NSCLC and MM, although it is noteworthy that the number of MM cases was rather small in this study. All missense mutations were predicted as being damaging to the protein product by either or both of the in silico tools, with two exceptions: EPHB1 Arg368Gly and EPHB6 Arg300Pro were predicted as neutral. Five mutations were confirmed as somatic, as they were validated in tumorous sample, but were not present in the paired normal sample: EPHA3 Gln676His, EPHA10 Gly260Val and Leu472Met, EPHB1 Thr981Met, and EPHB6 Asp653fs. Moreover, seven mutations were validated in tumorous tissue (Study I Table 3). The somatic nature remains elusive in most cases due to the lack of normal material.

Table 3. Novel non-synonymous Eph mutations and rare variants (less than 2 % in 1000 Human Genomes Project[a,b]) in NSCLC and MM (Adapted from Studies I and III). One patient can carry more than one mutation.

Gene	AA mutation	dbSNP/ COSMIC ID	MAF[a]	No of patients in Study I	No of patients in Study III	Total no of patients
EPHA1	p.Ser301del			1		1
	p.Gly398Trp				1	1
	p.Asn432Asp	rs150196251	0.0011[b]	1		1
	p.Arg492Gln	rs11768549	0.0062	1		1
	p.Ala585Thr	rs34178823	0.0002	1		1
	p.Arg844fs (1bp del)			1		1
EPHA2	p.Glu157Lys				1	1
	p.Pro350Thr	rs11543934	0.0004	2	1	3
	p.Arg762Ser				1	1
	p.Arg876His	rs35903225	0.0078		2	2
EPHA3	p.Tyr278Asn				1	1
	p.Gln676His[c]			1		1
	p.Tyr659Phe			1		1
	p.Ala777Gly	rs34437982	0.0004	3	1	4
EPHA5	p.Gly209Arg	rs202165566	0.00007[b]	1		1
EPHA6	p.Trp18Arg				1	1
EPHA7	p.Arg724Met			1		1
EPHA8	p.Gly45Ser	rs45498698	0.0030	1		1
	p.Asn432Asp	rs150196251	0.0011	1		1
	p.Arg441Gln	rs146978261	0.0001[b]		1	1
	p.Arg474Trp			1		1
	p.Pro607His	rs144329757	0.0034	1		1
	p.Ala611Ser				1	1
	p.Arg884His	rs62618734	0.0018	1*		1*
	p.Arg971His	rs139543017	0.00002[b]	1		1
EPHA10	p.Pro70His	COSM341849			1	1
	p.Gly260Val[c]				1	1
	p.Leu472Met[c]				1	1
EPHB1	p.Ile55Val	rs201772314	0.0002	1		1
	p.Arg222Trp				1	1
	p.Glu335Lys				1*	1*
	p.Arg368Gly			1*		1*
	p.Arg470Trp	rs202048188	0.0002*		1*	1*
	p.Arg497Leu	rs201303544	0.0004	1		1
	p.Val562Ile	rs202034365	0.003	1*		1*
	p.Val619Ala	rs375222902	0.0002[b]	1		1
	p.Leu843Met				1*	1*
	p.Thr981Met[c]	rs56186270	0.0002		1	1

EPHB2	p.Val762Phe			1		1
	p.Ala783Val				1*	1*
	p.Met849fs (2bp del)			1		1
EPHB3	p.Val266fs (1bp del)			1		1
	p.Asp785Asn				1	1
EPHB4	p.Gly221Ser				1	1
EPHB6	p.Gly122Ser	rs8177173	0.0032	3		3
	p.Gly272(stop)			1		1
	p.Arg300Pro			1		1
	p.Asp653fs[c] (1bp del)				1	1
	p.Glu711Lys	rs746719988	0.00002[b]	1		1

AA, amino acid

[a] Minor allele frequency detected in 1000 Human Genomes project, if not otherwise specified

[b] Minor allele frequency detected by Exome Aggregation Consortium (ExAC, Cambridge, MA, USA)

[c] Mutation validated as somatic; present in tumorous tissue, but not in paired normal tissue

* Malignant mesothelioma

Of the novel mutations, 33 % (9/27) occurred in the protein kinase domain, and two in the ligand binding domain of the receptors. Mutations which were predicted to be damaging in these functionally essential domains may cause deregulated activation, and changes in ligand binding, and thus promote possible oncogenic properties. However, possible oncogenic effect of the mutations located at other domains cannot be excluded. In the case of EPHA3, many mutations have been demonstrated both *in vitro* and *in vivo* to reduce kinase activity, ligand binding and downregulate the localization of the receptor on the cell surface (Lisabeth et al., 2012; Lahtela et al., 2013). EPHA3 has been described as a putative tumor suppressor (Zhuang et al., 2012; Lahtela et al., 2013), but in a later comprehensive study, the essentiality of EPHA3 in tumorigenesis of murine lung ADC driven by KRAS or TP53 alterations, was not confirmed (Lahtela et al., 2015).

In particular, mutations in EPHA3 and EPHA5 have been detected frequently in NSCLC (Davies et al., 2005; Ding et al., 2008; Imilienski et al., 2012; Saintigny et al., 2012), but mutations can be found in any of the Eph family genes (Imilienski et al., 2012; COSMIC) and are commonly spread along the gene (Saintigny et al., 2012; COSMIC). In this study, the tendency of EPHA3 and EPHA5 to gain mutations could not be confirmed. There is evidence that the Eph family mutations, such as EPHA2 mutations have been linked to cancer cell invasion and survival, and increased activation of mTOR signaling in SCC cells (Faoro et al., 2010). EPHA2 was recently identified as a tumor suppressor in KRAS positive ADC, i.e. there was inhibition of the RAS pathway and prevention of tumor cell proliferation after EPHA2 activation (Yeddula et al., 2015). Prognostic evidence of frameshift mutations of EPHB2 and EPHB4 has been found in colorectal cancer, as the mutations were associated with poorer prognosis (Alazzouzi et al., 2005). The exact role of Eph mutations remains still largely elusive, although somewhat more is known about altered expression statuses of Ephs, also in lung cancer. In NSCLC, elevated expressions of EPHA1, EPHA4 (Saintigny et al., 2012), EPHA5 and EPHA7 have been linked to better prognosis (Giaginis et al., 2014), whereas that of EPHA2 with a poorer prognosis (Brannan

et al., 2009). EPHA4 expression has also been associated with early stages of cancer and inflammation (Giaginis et al., 2014), and in another study, it was found to suppress cell invasion (Saintigny et al., 2012). Additional tumor suppressive roles have been detected in EPHB6, as it decreased metastasis and was silenced by promoter hypermethylation in NSCLC (Yu et al., 2010). On the contrary, elevated EPHB3 expression has been associated with increased cell growth and metastasis (Ji et al., 2011). Similarly, increased expressions of EPHA1, EPHA5 and EPHA7 have been associated with tumor cell proliferation (Giaginis et al., 2014). Moreover, EPHA7 expression have been linked to older age, fibrosis and smaller tumors (Giaginis et al., 2014), and EPHA2 expression to smoking history, EGFR activation (but no mutations) and KRAS mutations (Brannan et al., 2009). It has been postulated that Ephs may have different roles in tumorigenesis, and that they may even change from oncogenes to tumor suppressors during the course of tumorigenesis (Saintigny et al., 2012).

The rare variant and/or novel mutations in Ephs detected in four MM patients involved EPHB1; three with a novel mutation, and one with a rare variant (Study III Table 4; Table 3). EPHB1 mutations in MM have not been described previously. According to the COSMIC database, only one synonymous mutation of EPHB1 has been found in mesothelioma (Guo et al., 2015). Similar to the situation with NSCLC, more is known about the expression level of Ephs than the role of mutations in MM. Overexpression of EPHB2 has been found to associate with tumorigenesis of MM (Goparaju et al., 2013). In addition, the overexpressions of EPHA2 (Nasreen et al., 2006) and EPHB4 (Liu et al., 2013) have been observed in MM. Previously, inhibition of EPHA2 reduced cell growth and migration in MM cells (Nasreen et al., 2006). Intriguingly, Ephrin A1 ligand induced activation and phosphorylation of EPHA2 receptor and subsequently decreased EPHA2 expression and inhibited cell growth via ERK signaling in MM cells (Nasreen et al., 2007). A more detailed mechanism was revealed in a later study: Ephrin A1 activation causes let-7 miRNA expression, which suppresses the RAS oncogene and consequently tumor growth (Khodayari et al., 2011). This finding may be therapeutically interesting since it is known that certain tumors overexpress EPHA2 (Lee et al., 2013). In addition, inhibition of EPHB4 was found to suppress tumor growth, especially in combination with bevacizumab (anti-VEGF), in MM xenografts (Liu et al., 2013). These findings support the potential of Ephs to serve as therapeutic targets.

Eph mutations were frequently found simultaneously with other mutations in driver genes, such as *KRAS*, *EGFR*, *PI3K*, and there was also one concurrent case with the *ALK-EML4* fusion gene (Study I Supplemental Table S2; Study III Table 3). The prevalence together with the driver mutations suggests that they have more of a passenger role, although other possibilities e.g. enhancing driver capabilities of other alterations, cannot be excluded due to the involvement of Ephs in so many signaling pathways. For instance, EPHA3 mutations co-occurred with TP53 mutations statistically significantly in ADC, and some trend of concurrence was observed with KRAS mutations (Lahtela et al., 2015).

In this study, prognostic significance was not studied. However, two of those patients harboring a concomitant activating EGFR mutation and Eph mutation, were treated with EGFR-TKI and in both cases the disease continued to progress and survival was 15–18 months (Study I Supplemental Table S2). One patient harboring EPHA1 Arg844fs deletion and EGFR Leu858Arg, did show some response to treatment with gefitinib, but the disease

progressed. Another patient harboring EPHA5 G209R and EGFR exon 19 deletion was treated with gefitinib (after a relapse with erlotinib), and after a moderate response, the disease progressed.

In conclusion, Ephs clearly represent a significantly mutated receptor family in lung cancer and due to their diverse roles in modulating cell function especially their potential oncogenic and tumor suppressive roles, they can be viewed as a very intriguing group from a prognostic and therapeutic point of view. The studies on mutation and expression statuses of Ephs in the combination of other driver alterations should be continued.

5.2 EGFR mutations in non-small cell lung cancer (II, IV)

EGFR mutations are clinically significant in NSCLCs, and nowadays mutation testing is a routine part of diagnostics. Those patients harboring certain activating EGFR mutations are likely to respond to treatment with EGFR-TKIs.

In Study II, the EGFR mutation detection was performed with real-time PCR kits. The Therascreen EGFR PCR Kit produced valid test results as follows: 100 % (429/429) of the biopsy DNA samples, 95 % (37/39) of the CNB DNA samples, and 56 % (18/32) of the FNA DNA samples. Similarly, cobas EGFR Mutation Test produced the following valid results: 100 % (25/25) of the biopsy DNA samples, and 67 % (2/3) of the FNA DNA samples. Failed samples included FNA and CNB specimens, from which it is difficult to isolate DNA due to the small amount of tissue. In Study IV, mutations were investigated by targeted NGS on the Ion Torrent platform. Valid data was obtained from 425 out of 442 samples. This will be discussed in more detail in section of 5.4.

In summary, EGFR mutations were found in 11 % (58/510) of the patients, in Study II, and in 8 % (34/425) in Study IV. When combining the data, a total of 655 NSCLC samples were tested for EGFR mutations giving rise to a mutation frequency of 11.1 % (73/655). This is similar to, or close to, the highest frequencies, reported in Western or Central European patient cohorts (4.9–13.8 %) (Smits et al., 2012; Boch et al., 2013; Gahr et al., 2013; Ramlau et al., 2015; Skov et al., 2015).

In Study II, all mutated cases were diagnosed with ADC (Study II Table 1), and the new IASLC/ATS/ERS (2011) classification was used for the EGFR positive tumors. Eligible tumor material was available from 54 out of the 58 cases in which it was possible to determine ADC subgroup reliably: 98 % (53/54) were invasive ADCs and 74 % (40/54) were acinar predominant. With respect to the samples with acinar predominance, a mixed type was found in 18 samples: ten with papillary, of which one also displayed a micropapillary pattern and one with all four patterns; four with lepidic; three with a solid pattern; and one with non-mucinous minimally invasive ADC (MIA). Other predominance patterns of invasive ADCs were: lepidic (9 %, 5/54), solid (7 %, 4/54), micropapillary (4 %, 2/54) and papillary (4 %, 2/54). One EGFR positive ADC was classified as non-mucinous minimally invasive ADC (MIA).

Other investigators have also addressed this issue; Russell et al. (2013) found acinar and micropapillary subgroups to harbor most commonly EGFR mutations, however, there are many controversies about the association between EGFR mutations and ADC subgroups. Some recent studies have suggested that there may be a micropapillary subgroup (Li et al., 2013; Sun et al., 2014) and another lepidic subgroup which are most predominant in EGFR-mutant ADCs (Li et al., 2013; Villa et al., 2014). The variation may plausibly be due to differences in ethnic background, gender, smoking status and other features. Thus, no firm conclusions can be drawn, although it does seem that the EGFR mutation status does not correlate with any ADC subgroups. The classification is prognostically important (Russell et al., 2013; Sun et al., 2014; Xu et al., 2015).

In Study IV, four EGFR mutation positive cases were diagnosed with non-ADC histology: two SCCs (Leu858Arg and Val774Met), one LCC (Gly719Ala), and one NOS NSCLC (Ala767Val). Although EGFR mutations are strongly associated with ADC

histology (Dearden et al., 2013), they are occasionally found in other histologies (Dearden et al., 2013; Gahr et al., 2013; Ramlau et al., 2015; Skov et al., 2015).

When comparing the patients' other clinicopathological data and EGFR mutation status, in both studies, statistically significant associations were found between EGFR mutations and female gender ($p<0.001$) and never/light-smokers ($p<0.001$). However, mutations were also detected in smokers. With respect to the EGFR positive patients (n=73), 33 % (n=24) were never-smokers, 19 % (n=14) ex-light smokers, 29 % (n=21) ex-medium smokers and 19 % (n=14) current smokers. EGFR mutations and their higher prevalence in never-smoker and female patients have been observed from the very first EGFR mutation studies (Shigematsu et al., 2005) to more recent systematic reviews (Dearden et al., 2013; Midha et al., 2015). Nevertheless, the mutations occur also among (ex-)smokers (Dearden et al., 2013; Midha et al., 2015), although it has been claimed that a lower mutation frequency is likely in those subjects with a higher tobacco consumption status (Li et al., 2013). In addition to those well-known characteristics, in Study II, the mean age of patients with EGFR mutations was higher compared than that of the patients with wild-type EGFR ($p=0.001$).

Of the mutations detected, deletions in exon 19 and missense mutations Leu858Arg in exon 21 (as single mutations) were the most frequent, accounting for 71 % (52/73) of the EGFR mutation positive cases (Figure 7). This is very similar to the frequency that has been reported in previous studies on European patients (59–85 %) (Smits et al., 2012; Boch et al., 2013; Gahr et al., 2013; Skov et al., 2015). Compound EGFR mutants were found in seven cases in the following combinations: Gly719Ala+Ser768Ile in two cases, Gly719X+exon 20 insertion in two cases, Leu858Arg+Ser768Ile in one case, Leu858Arg+Asp761Tyr in one case and Leu858Arg+Ala871Thr+Gly735Ser in one case. Compound mutants have been reported in many studies, and their proportion of EGFR mutated cases has been estimated as between 5 to 25 % (Hsieh et al., 2006; Hata et al., 2010; Kobayashi et al., 2013; Keam et al., 2014; Stone et al., 2014; Peng et al., 2015; Kim et al., 2016). Their possible clinical significance remains obscure. They have been linked to poorer prognosis (Kim et al., 2016), but conflicting observations have been made and it is not clear whether those patients with a classical activating mutation combined with a rare mutation exhibit a similar response to treatment with EGFR-TKIs compared to the patients with classical activating mutations alone (Kobayashi et al., 2013; Keam et al., 2014; Peng et al., 2015). Some studies have shown that the differences in response do seem to depend on the complex mutation type (Hsieh et al., 2006; Hata et al., 2010; Wu et al., 2011a).

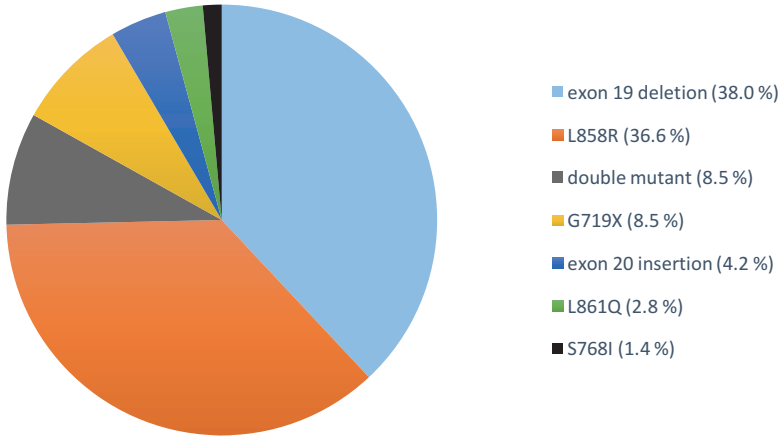

Figure 7. Proportional distribution of EGFR mutations in mutation positive NSCLC patients (n=71). Rare mutations included Val774Met and Ala767Val.

In Study II, EGFR mutations were detected in 11 % of the asbestos-exposed patients vs. 10 % in non-exposed group (p=0.498). In Study IV, no EGFR mutations were seen in asbestos-exposed group, but the number of the exposed patients was rather small (n=29) and statistical significance was not found (p=0.146). This suggests that EGFR mutations do not have any association with asbestos-exposure, or may be slightly more prevalent in non-exposed patients. This is concordant with a previous study by Andujar and colleagues (2013), who did not observe any difference in the EGFR mutation prevalence between asbestos-exposed (2/50) and non-exposed NSCLC (6/50).

In Study II, the association between EGFR mutation status and chronic obstructive pulmonary disease (COPD) was also assessed. EGFR mutations were seen more often in the patients without COPD: 13 % (51/399) vs. 6 % (7/111), suggesting that EGFR mutations occur less frequently among COPD patients (p=0.057). This finding is in good agreement with other studies, where the prevalence of EGFR mutations has been higher in non-COPD patients, varying between 15–37 %, compared to 5–16 % in COPD patients (Suzuki et al., 2010; Lim et al., 2015; Saber et al., 2016).

In Study II, the median follow-up time was 23 months (range 0–112). Overall, data from 65 % of the patients was eligible and suitable for survival analysis. There was no statistical difference detected in OS between ADC patients with wild-type EGFR (68.6 months; 95 % CI, 62.6–74.7) and ADC patients with the mutated EGFR (52.2 months; 95 % CI, 40.2–64.2, p=0.460) (Study II Figure I and Table 3). Similarly, no statistical differences were seen in OS between EGFR mutation positive and negative patients when only non-smoker ADC patients were analyzed. Whether or not EGFR mutation is a prognostic marker remains controversial. In a recent systemic review of 22 studies and a meta-analysis evaluating 16 reports of early-stage resected NSCLC, no statistically significant prognostic role of EGFR mutations was found in OS (p=0.12) or PFS rates (p=0.65) (Zhang et al., 2014). This supports our finding, although the follow-up time was rather short and a large number of patients were still alive at the termination of this study.

Clinically significant EGFR mutations have been commonly reported to be mutually exclusive with other driver alterations, such as ALK-EML4 fusion, BRAF Val600Glu and KRAS mutations (Dearden et al., 2013; Gainor et al., 2013; Tissot et al., 2016). EGFR mutations were found to be significantly mutually exclusive with KRAS mutations as well as with the ALK-EML4 fusion in ADC patients (Dearden et al., 2013). This was also seen in this study with one exception. In Study IV, we identified one patient with a total of 15 hot spot mutations including EGFR Leu858Arg and KRAS Gly13Cys. Concurrences of driver alterations are seen occasionally (Ulivi et al., 2015; Lee et al., 2016; Sahnane et al., 2016), and for instance, a concurrent KRAS mutation with the ALK fusion has been proposed to be a marker of a poor prognosis (Ulivi et al., 2015).

These two separate analyses on partly overlapping material also provided an opportunity to investigate the concordance of the results obtained by real-time PCR and targeted NGS. Briefly, a concordance of 99.7 % was obtained when considering only the mutations detectable by the PCR kit. Only one discordant case, i.e. a Gly719X mutation was detected in one sample by PCR, but not by NGS. However, this produced only a weak signal in PCR and was assessed as being unclear. We noted other advantages with NGS i.e. two insertions in exon 20 which were not included in the PCR panel were detected by NGS. These results confirm NGS to be highly sensitive method, and at least comparable to the so-called golden standard PCR methods. Results of this comparison are published in Mäki-Nevala et al. (2016), and are in agreement with previous studies (Belardinilli et al., 2015; Malapelle et al., 2015; Fujita et al., 2015).

5.3 Asbestos-associated mutations in lung adenocarcinoma and malignant mesothelioma (III)

Asbestos is an important occupational carcinogen and a major cause of MM, a very aggressive form of cancer with a dismal prognosis. It should be noted that lung cancer may also be caused by exposure to asbestos. Although the molecular features are rather well established in ADC and MM, it does seem that asbestos-exposure related alterations have remained obscure.

We performed exome sequencing on 47 tumor specimens including 26 ADC and 21 MM (19 pleural and two peritoneal) cases, and nine paired normal samples of ADC patients. Exome sequencing could be performed successfully on all samples. The mean target coverage was 38 (range 13–54), for FFPE samples it was 37 (range 13–54) and for FF samples 39 (range 20–54). The exome data were filtered through various steps, and eventually eight genes were identified as possibly being associated with asbestos-exposure: *BAP1*, *COPG1*, *INPP4A*, *MBD1*, *SDK1*, *SEMA5B*, *TTLL6* and *XAB2* (Study III Table 2). These genes were exclusively mutated in more than two samples of asbestos-exposed patients, and harbored novel variants which were predicted as damaging, and not reported in dbSNP or any paired normal sample.

BAP1 mutations, and SDK1 Gln963Ter were validated as somatic by amplicon sequencing, as they were present in tumorous tissue, but not in paired normal tissue material. For one BAP1 mutation positive case, paired normal tissue was not available, but the mutation was validated in tumor material. Moreover, the following mutations were reproducible in tumor tissue: COPG1 Cys230Arg, INPP4A Lys954Arg, SEMA5B Thr1040Pro and TTLL6 Glu56fs. The MRPL1 Tyr87Cys mutation was not found in one paired normal sample supporting its somatic nature, but unfortunately sequencing was not successful in the corresponding tumor tissue material. These genes with some supportive evidence for their somatic and/or true positive nature will be discussed in more detail below.

A recurrent novel mutation was found in mitochondrial ribosomal protein L1 (MRPL1) (Tyr87Cys) occurring exclusively in three asbestos-exposed patients, in two MM and one ADC. Six MRPL1 mutations confirmed as being somatic have been reported previously; four in ADC and two in SCLC histology (COSMIC, 08/05/2016). No mutations have been reported to occur in pleura. MRPL1 is a nuclear gene encoding a 39S subunit of the mitochondrial ribosome (mt-rRNA) and is thus involved in peptidyl transferase activity in mitochondrial protein synthesis (Elson et al., 2015). Its role in cancer has not been described, and it can only be speculated that mutations may lead to dysfunction of some mitochondrial protein translational event. The mitochondrial genes encode proteins involved in cell metabolism, and these kinds of alterations are one hallmark of cancer (Weinberg and Hanahan, 2011).

ROS species derive from cell metabolism and are especially intriguing in asbestos-related cancer, as asbestos is known to induce inflammation in the lungs and trigger the production of ROS species (reviewed in Sekido, 2013). With respect to the mutated genes, also *sidekick cell adhesion molecule 1* (*SDK1*) and *inositol polyphosphate-4-phosphatase type I A* (*INPP4A*) have been related to oxidative stress. INPP4A is involved in dephosphorylation of second messenger molecules (inositols) that regulate various signaling pathways. For instance, the PI3K/AKT pathway, which is commonly altered in

lung cancer, and phagocytosis are negatively regulated by INPP4A (Ivetac et al., 2009; Nigorikawa et al., 2015). Activation of PI3K/AKT pathway can promote oxidative stress (Kim et al., 2005). Phagocytosis is also an event linked to molecular pathology of MM, as macrophages engulf, but do not seem capable of disposing of asbestos fibers and this promotes increased production of ROS species (Sekido, 2013). *INPP4A* has been predicted as a candidate gene in asthma, and some polymorphisms in this gene plausibly inhibit INPP4A function (Sharma et al., 2008). Moreover, the gene has been observed to serve as an important regulator in inflamed lungs of mice (Agrawal et al., 2009; Aich et al., 2012a and b).

SDK1 is a cell adhesion molecule that has been found to be expressed during the cellular stress induced by stress signals, particularly ROS species. Its expression is controlled by at least two transcription factors, signal transducer and activator of transcription 3 (STAT3) and nuclear factor-κB (NF-κB) (Yoon et al., 2012). SDK1 is located at common fragile site of the genome, which is prone to break in cancer (Bosco et al., 2010). More recently, somatic mutations of SDK1 have been detected in ADC (Imilienski et al., 2012), and in another study *SDK1* was predicted to be a putative driver gene in lung ADC, particularly if there is a mutation G399D (Zhao et al., 2014). Moreover, recurrent somatic mutations were detected in adrenocortical carcinoma (Juhlin et al., 2015), and most intriguingly SDK1 was predicted to be associated with an increased MM risk in Australian patients (Cadby et al., 2013). These findings suggest that SDK1 should receive more attention in the future, as it may be an important actor in tumorigenesis.

Mutations in BAP1 and COPG1 occurred only in MM patients, but as asbestos-non-exposed cases of MM are very rare, no firm conclusions about their association with asbestos-exposure can be drawn. BAP1 mutations are well-established in MM, both as germline (Testa et al., 2011) and somatic (Bott et al., 2011: Yoshikawa et al., 2012; Zauderer et al., 2013). BAP1 is a nuclear ubiquitin C-terminal hydrolase that interacts with multiple proteins, and thus is involved in many cellular functions, such as in chromatin modeling, DNA damage, cell cycle and growth (Eletr et al., 2011). The mutations detected in this study have not been reported previously in MM, but are located in a region which contains the ubiquitin carboxyl hydrolase (UCH) site of the protein, which is frequently found to be mutated in MM (Bott et al., 2011), supporting their pathological role.

Coatomer protein complex subunit gamma 1 (*COPG1*) encodes a subunit in a coatomer protein complex involved in vesicle transport in protein secretion (Hahn et al., 2000). Increased expression of alpha subunit of the same complex was observed in mesothelioma cells (but not in pleural cells), and its knockdown was associated with a tumor suppressive role and apoptosis (Sudo et al., 2010). Previously, COPG has been found to be highly expressed in both lung cancer cells and lung cancer-derived endothelial cells (Park et al., 2008). Mutations of subunits involved in this protein complex are not well-established, although some have been reported in COSMIC.

Semaphorin 5B (*SEMA5B*) encodes for a member of the semaphorin protein family; it regulates the development of the nervous system. Somatic mutations in *SEMA5B* have been reported in various tumors, also in lung and pleura (COSMIC), but their role is unknown. However, *SEMA5B* was proposed to be a candidate gene for susceptibility of esophageal carcinoma in a Chinese population (Wu et al., 2011b). Hirota et al. (2006) noted that SEMA5B knockdown decreased the renal cell carcinoma cell viability which was

interpreted as evidence for an oncogenic role. Multiple other members of semaphorins are associated with tumorigenic events, such as angiogenesis (reviewed in Worzfeld and Offermanns, 2014).

Tubulin tyrosine ligase-like 6 (*TTLL6*) encodes a protein involved in apoptosis, and that is overexpressed particularly in the testis (Chen et al., 2006). Somatic mutations similarly have been reported occasionally in lung tumors (COSMIC), but their possible role and the whole gene's role in cancer have remained elusive.

We studied some genes which can function as drivers and/or are known to be mutated in lung ADC and/or MM. Those were: *BRAF*, *EGFR*, *ERBB2*, *HRAS*, *KRAS*, *MET*, *NRAS, PIK3CA*, *STK11* for both ADC and MM, plus *BAP1*, *CUL1*, *CDKN2A*, *NF2* for MM. The mutations predicted as damaging were detected in Study III (Table 3): *BRAF* (only ADC), *EGFR*, *ERBB2* (only ADC), *KRAS* (only ADC), *MET* (only ADC), *STK11*, *BAP1* (only MM) and *NF2* (only MM). The most frequently mutated genes were *KRAS* (nine ADCs), *BAP1* (four MMs) and *STK11* (two ADCs and one MM). As a vast majority (25/26) of the ADC patients were smokers, KRAS mutations can be expected to occur frequently (Dearden et al., 2013). This may be the reason why we failed to find any sign of activating EGFR mutations. Mutations detected in these genes seem to occur in both asbestos-exposed and non-exposed samples.

5.4 Hot spot mutations in non-small cell lung cancer (IV)

Particular regions in the genome are more prone to harbor cancer-related somatic mutation, these kinds of regions are called mutational hot spots. We studied the hot spots of 22 lung cancer-related genes in 425 NSCLC patients. Study was conducted on FFPE samples by targeted NGS using a panel including the following genes: *AKT1, ALK, BRAF, CTNNB1, DDR2, EGFR, ERBB2, ERBB4, FBXW7, FGFR1, FGFR2, FGFR3, KRAS, MAP2K1, MET, NOTCH1, NRAS, PIK3CA, PTEN, SMAD4, STK11* and *TP53*. This screening study is the first to examine a Finnish NSCLC patient cohort. Due to the history of genetic isolation, and the fact that certain somatic mutations are associated with ethnic background, the Finnish NSCLC patients serve as an attractive study material.

The sequencing was performed on 442 tumor samples, data from 425 met the predefined quality criteria. The resulted samples had an average coverage of 1623 (range 146–4137). The samples discarded from the analysis included those with low sequencing coverage and/or libraries with low DNA concentration.

The mutations were divided into those reported in COSMIC (referred to as COSMIC mutations), and novel mutations including those not reported in dbSNP or COSMIC databases. COSMIC mutations were seen in 77 % of the patients. The majority of mutation positive patients (46 %) carried one mutation, 25 % had two mutations, and the others harbored from three up to 15 mutations. The most frequently mutated gene was *TP53* with a frequency of 46 %, as was expected (Dearden et al., 2013; Cancer Genome Research Atlas Network, 2012 and 2014). The second most commonly mutated gene was *KRAS* (26 %) and the third was *MET* (15 %) (Study IV Figure 1 and 2, and Table 2). No COSMIC mutations were detected in the following genes included in the panel: *ALK*, *DDR2*, *FGFR1*, *FGFR2*, *MAP2K1* and *SMAD4*. However, non-synonymous novel mutations predicted as damaging were detected in those genes (Study IV Supplemental File S1 and Figure 2a).

A total of 195 non-synonymous novel variants were detected and nearly half (n=93) were predicted to be damaging to the protein product (Study IV Supplemental Table S1). It is worthwhile highlighting some of these novel variants; they included some probable pathogenic variants, such as two EGFR exon 19 deletions, and one ERBB2 exon 20 frameshift insertion. ERBB2 exon 20 insertions are potentially sensitive to HER2 inhibitors (Mazieres et al., 2016). EGFR exon 19 deletions are sensitive to EGFR-TKIs, and they are believed to represent their own entity and differ from the Leu858Arg mutation positive NSCLCs (Lee et al., 2015).

The mutations and their incidence resembled largely that described in Western NSCLC cohort previously. Some differences can be pointed out, such as a high frequency of BRAF and MET mutations and a lower frequency of ERBB2 and STK11 mutations. BRAF mutations were found in 5.6 % of the patients, as the reported average frequency in NSCLC is approximately 3 % (Chen et al., 2014). However, the proportion of Val600Glu mutations was much lower in this study. It has been found to account for approximately a half of the BRAF mutants in NSCLC; in this study, the proportion of Val600Glu mutants was only 17 % (4/24). The total frequency of BRAF Val600Glu was only 0.9 %, the reported frequencies being around 2 % (Brustugun et al., 2014). BRAF mutants found in this study occurred in amino acid residues from 466 to 605 (Study IV Table 2). The most frequently mutated was codon 469 with nine mutated cases. The clinical significance of BRAF mutants in NSCLCs

remains unclear (Tissot et al., 2016), although the mutations have been shown to alter protein activity (Davies et al., 2002; De Falco et al., 2008; Dahlman et al., 2012).

Previously MET mutations have been reported in 7 % of ADCs (Cancer Genome Research Atlas Network, 2014), while in this study, these kinds of mutations were found in 16 %. The mutations (Asn375Ser and Thr1010Ile) found in this study are also present in dbSNP, so their somatic nature remains questionable. However, the Thr1010Ile mutation was detected at a higher frequency in this NSCLC cohort than has been reported for the Finnish population in 1000 Genomes (NCBI Genomes Browser) (6 % vs. 2 %). The association of Asn375Ser with lung cancer tumorigenesis or prognosis is disputable, one study finding its association with a shorter PFS (Cao et al., 2014), but this has not been confirmed (Shieh et al., 2013). Similarly, the Thr1010Ile mutation has been proposed to contribute to aggressive behavior in breast cancer (Liu et al., 2015) and lung cancer cells (Ma et al., 2003; Tengs et al., 2006), but conflicting results have been reported (Schmidt et al., 1997; Tyner et al., 2010). Thus, their role in lung cancer remains undetermined. In this study, their high frequency and concurrence with driver mutations suggests that they have a germline origin, although they may affect prognosis and tumorigenesis.

The concurrences of COSMIC mutations were also studied (Study IV Table 3; Table 4). TP53 mutations were very frequently detected together with other variants, also with pathogenic KRAS and EGFR mutations. There is no consensus on whether TP53 mutations are prone to occur with other mutations. In a recent meta-analysis, TP53 mutations were seen together with EGFR and KRAS in 7.3–7.8 % of the patients, respectively (Dearden et al., 2013). On the other hand, in a comprehensive molecular screening of lung ADC, TP53 mutations were more rarely seen with oncogenic EGFR, KRAS and BRAF (Cancer Genome Research Atlas Network, 2014). Previously, STK11 mutations have been found to be associated with KRAS mutations (Koivunen et al., 2008; Dearden et al., 2013; Pecuchet et al., 2015), and also with TP53 (Dearden et al., 2013), a tendency which was also seen in this study. The high frequency of mutated TP53 found here might be one reason for its higher concurrence with other gene mutations, and/or the fact that mostly early stage of NSCLCs were studied. It has been suggested that the stage of the disease may affect the prevalence of TP53 mutations, but it is far from clear whether TP53 mutations emerge at earlier or late stages of the disease (Rivlin et al., 2011). Barnett's et al. (1999) prospective study on esophageal ADC suggested that they occur at early stages in tumorigenesis. Moreover, TP53 is the most commonly mutated gene in many cancers, i.e. in approximately 50 % of all cancers (Hollstein et al., 1991), and similarly NSCLC (Dearden et al., 2013) carry mutated p53. Thus, TP53 mutations likely occur concurrently with other mutations, also with driver changes.

Table 4. The proportions (%) of concurrent COSMIC mutations of the mutated cases, calculated per row. Adapted from Study IV Table 3.

Gene	N	EGFR	ERBB2	ERBB4	FGFR3	MET	AKT1	PIK3CA	KRAS	NRAS	BRAF	CTNNB1	NOTCH1	FBWX7	PTEN	STK11	TP53
EGFR	34		0	3	0	12	0	3	6	0	3	6	0	0	0	0	53
ERBB2	3	0		0	0	33	0	0	33	0	0	0	0	0	0	0	33
ERBB4	2	50	0		0	0	0	50	100	0	50	100	0	0	0	0	100
FGFR3	2	0	0	0		0	0	0	0	0	0	0	0	0	0	0	100
MET	65	6	2	0	0		0	9	28	2	5	3	0	0	0	0	8
AKT1	1	0	0	0	0	0		0	0	0	0	0	0	0	0	0	100
PIK3CA	18	6	0	6	0	33	0		17	0	6	6	6	6	0	0	61
KRAS	111	2	1	2	0	16	0	3		0	2	5	1	1	0	2	35
NRAS	1	0	0	0	0	100	0	0	0		0	0	0	0	0	0	0
BRAF	24	4	0	4	0	13	0	4	8	0		8	4	0	0	0	50
CTNNB1	10	20	0	20	0	20	0	10	50	0	20		0	0	10	0	60
NOTCH1	3	0	0	0	0	0	0	33	33	0	33	0		0	0	0	67
FBXW7	2	0	0	0	0	0	0	50	50	0	0	0	0		0	0	0
PTEN	5	0	0	0	0	0	0	0	0	0	0	20	0	0		0	100
STK11	8	0	0	0	0	0	0	0	25	0	0	0	0	0	0		38
TP53	197	9	1	1	1	13	1	6	20	0	6	3	1	0	3	2	

Light red: 0–5 %; yellow: 6–9 %; orange: 10–20 %; light green: 21–49 %; green: 50 % and more

There were also other COSMIC mutations in addition to EGFR, e.g. statistically significant associations were found between PIK3CA ($p=0.016$) and TP53 ($p<0.001$) mutations and SCC histology, and *KRAS* mutations and ADC ($p<0.001$). The CTNNB1 mutations seemed to be more likely to be present in light ex-smokers ($p=0.011$) (Study IV Table 4; Table 5). The associations between PIK3CA and TP53 with SCCs are well-established (Dearden et al., 2013; Li et al., 2014b), as is the link between KRAS mutations and ADC (Dearden et al., 2013). The significance of smoking status and its possible association with mutations remain to be clarified. In this study, no mutation other than EGFR and CTNNB1 was found to be significantly associated with smoking status, the latter has not been described before. Although statistical significance was found, the number of CTNNB1 positive cases was still rather small. There were some tendencies, such that TP53 mutations would be more common among smokers, which was also an observation in a recent meta-analyses on lung

cancer (Liu et al., 2014) and esophageal cancer (Wu et al., 2015). KRAS mutations have been found more often in smokers (Dogan et al., 2012), but in this study, no clear association was detected, although the mutation incidence was somewhat higher among ever-smokers compared to never-smokers (27 % vs. 18 %, p=0.334).

No association was found between the likelihood of asbestos-exposure and COSMIC mutations. However, some tendency was detected with respect to the TP53 and EGFR mutations, the latter has been discussed in section 5.2. TP53 was more frequently mutated among the asbestos-exposed group (46 % vs. 62 %, p=0.118). In the previous study of Andujar et al. (2013), no statistically significant difference was found in the prevalence of EGFR, KRAS or TP53 mutations between asbestos-exposed and non-exposed NSCLCs. However, some TP53 polymorphisms were associated with asbestos-exposed patients (Andujar et al., 2013). In some earlier studies, TP53 mutations have been reported to be more prevalent in asbestos-exposed patients (Wang et al., 1995), although in another study, no statistical difference could be seen (Husgafvel-Pursiainen et al., 1999). Other clinical factors, as well as the method used to determine asbestos-exposure, and less-sensitive mutation testing may play a role, and introduce variation into those studies which have usually been conducted on a relatively small numbers of cases.

Due to the short follow-up time and the fact that many of our patients were alive at the end of the follow-up, it was not deemed feasible to conduct a survival analysis, but some observations can be made. Patients with multiple mutations were observed to die more frequently due to lung cancer compared to patients carrying fewer mutations (53 % of the patients with more than three mutations vs. 20–37 % of patients with two, one or no mutations). Bria et al. (2015) have also reported the association of multiple mutations and poorer survival. In this respect, the mutation burden may be both a predictive and prognostic marker.

Table 5. Genes with COSMIC mutations. Mutation frequencies presented in descending order, and associations with clinical data.

Gene	Mutation frequency (% (N))	Statistically significant associations with clinical features (p<0.05)	Tendencies to clinical features/patient characteristics (0.05<p<0.15)
TP53	46.4 (197/425)	SCC (p<0.001)	Tendency to HS (p=0.135) Tendency to ES (p=0.061) Tendency to AE (p=0.118)
KRAS	26.1 (111/425)	ADC (p<0.001)	
MET	15.3 (65/425)		
EGFR [a]	11.1 (73/655)	ADC (p<0.001/0.037) NS/LS (p<0.001) Female (p<0.001)	Tendency to ANE (p=0.146)[b]
BRAF	5.6 (24/425)		Tendency to ADC (p=0.080) Tendency to NS (p=0.146) Tendency to female (p=0.145)
PIK3CA	4.2 (18/425)	SCC (p=0.016)	
CTNNB1	2.4 (10/425)	LS (p=0.011)	
STK11	1.9 (8/425)		Tendency to ES (p=0.097) Tendency to male (p=0.085)
PTEN	1.2 (5/425)		Tendency to male (p=0.071)
ERBB2	0.7 (3/425)		
NOTCH1	0.7 (3/425)		
ERBB4	0.5 (2/425)		
FGFR3	0.5 (2/425)		
FBWX7	0.5 (2/425)		
AKT1	0.2 (1/425)		
NRAS	0.2 (1/425)		

AE, asbestos-exposed; ANE, asbestos non-exposed; ES, ever-smoker; LS, light ex-smoker; NS, never-smoker

[a] Combined results of Study II and Study IV

[b] Association found in Study IV, not in Study II

6 CONCLUSIONS AND PROSPECTS

In this study, a large series of more than 400 Finnish NSCLC and MM patients were studied for molecular genetic markers by estimating their frequencies, concurrences and associations with clinical data from the patients. The principal methods used in this study were NGS, and PCR-based mutation testing.

In Studies I and III, Eph genes were often found to be mutated, novel mutations were found with a frequency of 18 %. Mutations were seen in both lung ADC and MM patients. All MM cases harboring either a mutation or a rare variant of Eph, carried the mutation/variant in EPHB1. Eph mutations did not associate with clinical features, such as smoking status, gender or asbestos-exposure. Ephs are an intriguing group of proteins from a therapeutic point of view, as they are involved in a complex network of signaling pathways and play diverse roles in many crucial cellular functions, such as in cellular adhesion and angiogenesis. Moreover, due to the associations between their altered expression and pathogenic features and/or prognosis of the patients, they may serve as novel biomarkers for patient management. Thus, they may be a crucial component of the tumorigenesis process and could be targeted by drugs. As confirmed here, they are frequently mutated in lung cancer specimens; clearly more needs to be known about these alterations as well as how these changes alter cellular functions.

EGFR mutations were found in 11 % of the NSCLC patients. These mutations were significantly associated with never- or light-smoking status, female gender and adenocarcinoma histology, in accordance with previous studies. However, mutations occurred occasionally also in non-ADC histological subgroups and even in current smokers. Thus, this study also strongly suggests that histological subgroup or smoking status should not be an indication for determining whether or not the subject should undergo EGFR mutation testing. In addition, multiple compound EGFR mutants were found, accounting for 10 % of mutation positive NSCLCs.

Study III on asbestos-exposure associated mutations was one of the first studies indicating that there seemed to be some molecular differences, also at the level of the DNA bases, between asbestos-exposed and non-exposed lung cancer patients. It appears that all of the asbestos-exposed patients harbored mutated genes related to oxidative stress, which is one outcome of asbestos-exposure. However, known driver mutations, such as those in KRAS, do not seem to associate with asbestos exposure. Nevertheless, some genetic alterations in asbestos-exposed lung cancer are apparently different from those present in non-exposed individuals, and by clarifying their nature, it may be possible to devise novel therapy strategies against these very aggressive mesothelioma tumors.

The mutations in the genomic regions that are prone to cancer-related mutations (hot spots), seem to occur in Finnish NSCLC patients similarly as those in their counterparts in other Western countries. Some differences were observed, such as a higher overall incidence of BRAF mutations, but a lower incidence of the BRAF Val600Glu mutation. The most frequently mutated genes were TP53 (46 %) and KRAS (26 %). The concurrence of TP53 mutations with other hot spot mutations was very evident. A total of 77 % of the NSCLC patients carried one or more hot spot mutations.

Overall, this thesis provides comprehensive knowledge of the mutation spectrum present in Finnish NSCLC patients. The novel mutations found in this study provide material for

future studies on molecular investigation of lung cancer since they may represent important markers for tumorigenesis and it would be beneficial to clarify their associations and concurrences since this could help to develop novel targeted therapies. The studies revealed the benefits of NGS technology for investigating cancer-related mutations in archival FFPE material. The great advantage of NGS is that it is possible to examine multiple mutations, both novel and known, in great detail and in different genes simultaneously during a single test run. NGS was also found to be at least as sensitive as the golden standard PCR methods.

In the future, one major focus will be on developing combined targeted therapies, based on predictive markers and comprehensive molecular testing. NGS methods are gradually replacing the standard methods in cancer diagnostics. It is important to reveal NSCLC patients' molecular characteristics in addition to the histological and morphological features of their tumors if one wishes to deliver the optimal therapy with the best prognosis. A knowledge of concurrences of the alterations and in particular, possible resistance mechanisms, are crucial since at present, many of the newly developed drugs show initial efficacy but this wanes with the emergence of resistance.

ACKNOWLEDGEMENTS

This thesis was carried out in the Department of Pathology, Faculty of Medicine, University of Helsinki, during 2013–2016. I thank the Director of Medicum, Professor Tom Böhling, and the Head of Department of Pathology, Professor Veli-Pekka Lehto, for enabling this work to be completed by providing research facilities. I am grateful for the financial support received from the Cancer Society of Finland, the Sigrid Jusélius Foundation, the special governmental subsidy research funds provided to the Helsinki and Uusimaa Hospital District (HUS EVO), the Finnish Work Environment Fund, Pfizer Oy, AstraZeneca AS, the Finnish Culture Foundation and the University of Helsinki (dissertation completion grant).

I express my sincere gratitude to everyone who made this work possible, especially the following persons:

My supervisors, Professor Sakari Knuutila and Dr. Virinder Kaur Sarhadi, for all valuable support. First, I am grateful to Professor Knuutila for giving me this opportunity to be part of his friendly research group, and for introducing me to academic research and the field of cancer genetics. I am grateful that he gave me this responsibility and allowed me to gain my independence to learn; these are invaluable qualities for a student and researcher. I want to express my gratitude to both supervisors for the inspirational and instructive discussions, it has been a privilege to work under your positive and encouraging guidance.

The thesis committee members Dr. Sippy Kaur and Dr. Suvi Savola for their support and instructive role.

The official pre-examiners, Professor Vesa Kataja and Docent Henna Tyynismaa, for their critical review and valuable comments and editing that improved the thesis manuscript.

The opponent, Docent Paavo Pääkkö, for accepting the invitation to be the official opponent at the thesis defense.

Dr. Ewen MacDonald for editing the language of the final manuscript.

The coauthors of the studies, Zoe Dawson, Maria Gomez, Pekka Ellonen, Kirsti Husgafvel-Pursiainen, Eeva Kettunen, Aija Knuuttila, Sonja Lagström, Mike Morel, Mikko Rönty, Aino Telaranta-Keerie and Henrik Wolff for their collaboration via their scientific and intellectual contributions to this work. Aija Knuuttila is especially deserving of thanks for her indispensable clinical contribution and help with collection of the patients' clinical data.

My colleagues, former and present members of the CMG group, especially Gemma Armengol, Leo Lahti, Lauri Lehtimäki, Neda Mosakhani, Tarja Nieminen, Noora Porkka, Ilari Scheinin, Katja Tuononen, Aino Wirtanen, Omar Youssef, and all of the others, for their help and for creating such a friendly working environment. I would especially like to thank Milja Tikkanen for helping me with so many practical issues and for becoming such a good friend. Tiina Wirtanen is especially deserving of thanks for her thorough guidance on how to do laboratory work at the beginning of my PhD studies.

My lovely friends for their great company, sharing all the good moments and helping me through the difficult ones. I would like to express my gratitude especially to Paula and Maria, without whom I would not be what I am now. I also thank Tytti, my housemate and good friend for providing good company during the more stressful times. Carlo, my love, I am so grateful you became part of my life, thank you for all the support and happiness you

bring to me. I am grateful to my sister Saija and her husband Jarkko for all of the support, good company and great conversations, which often added a new perspective to my life. My deepest gratitude goes to my mother, thank you for your endless love, care and support during my whole life, and for having such strong faith in me all these years. I only wish I could express my deepest gratitude also to my father, whom we lost far too early, how I wish he could see where I am now and who I have become.

Helsinki, 2016

Satu Mäki-Nevala

WEB-BASED RESOURCES

1000 Human Genomes Project	http://www.1000genomes.org/home
BCFtools	http://samtools.github.io/bcftools
COSMIC	http://cancer.sanger.ac.uk
ExAC	http://exac.broadinstitute.org
GeneCards	http://www.genecards.org
National Human Genome Research Institute	http://www.genome.gov
NCBI dbSNP	http://www.ncbi.nlm.nih.gov/SNP
PROVEAN	http://provean.jcvi.org
SeqPrep	https://github.com/jstjohn/SeqPrep
SIFT	https://sift.jcvi.org

REFERENCES

Abe S, Morita Y, Kaneko MK, Hanibuchi M, Tsujimoto Y, Goto H, Kakiuchi S, Aono Y, Huang J, Sato S, Kishuku M, Taniguchi Y, Azuma M, Kawazoe K, Sekido Y, Yano S, Akiyama S, Sone S, Minakuchi K, Kato Y and Nishioka Y. A novel targeting therapy of malignant mesothelioma using anti-podoplanin antibody. J Immunol. 2013; 190:6239-6249.

Abramson R. Overview of Targeted Therapies for Cancer. *My cancer genome.* https://www.mycancergenome.org/content/molecular-medicine/overview-of-targeted-therapies-for-cancer . Accessed 05/13. 2016.

Agrawal A, Sinha A, Ahmad T, Aich J, Singh P, Sharma A and Ghosh B. Maladaptation of critical cellular functions in asthma: Bioinformatic analysis. Physiol Genomics. 2009; 40:1-7.

Aich J, Mabalirajan U, Ahmad T, Agrawal A and Ghosh B. Loss-of-function of inositol polyphosphate-4-phosphatase reversibly increases the severity of allergic airway inflammation. Nat Commun. 2012a; 3:877.

Aich J, Mabalirajan U, Ahmad T, Khanna K, Rehman R, Agrawal A and Ghosh B. Resveratrol attenuates experimental allergic asthma in mice by restoring inositol polyphosphate 4 phosphatase (INPP4A). Int Immunopharmacol. 2012b; 14:438-443.

Alazzouzi H, Davalos V, Kokko A, Domingo E, Woerner SM, Wilson AJ, Konrad L, Laiho P, Espin E, Armengol M, Imai K, Yamamoto H, Mariadason JM, Gebert JF, Aaltonen LA, Schwartz S,Jr and Arango D. Mechanisms of inactivation of the receptor tyrosine kinase EPHB2 in colorectal tumors. Cancer Res. 2005; 65:10170-10173.

Andujar P, Pairon JC, Renier A, Descatha A, Hysi I, Abd-Alsamad I, Billon-Galland MA, Blons H, Clin B, Danel C, Debrosse D, Galateau-Salle F, Housset B, Laurent-Puig P, Le Pimpec-Barthes F, Letourneux M, Monnet I, Regnard JF, Validire P, Zucman-Rossi J, Jaurand MC and Jean D. Differential mutation profiles and similar intronic TP53 polymorphisms in asbestos-related lung cancer and pleural mesothelioma. Mutagenesis. 2013; 28:323-331.

Arcila ME, Oxnard GR, Nafa K, Riely GJ, Solomon SB, Zakowski MF, Kris MG, Pao W, Miller VA and Ladanyi M. Rebiopsy of lung cancer patients with acquired resistance to EGFR inhibitors and enhanced detection of the T790M mutation using a locked nucleic acid-based assay. Clin Cancer Res. 2011; 17:1169-1180.

Arcila ME, Nafa K, Chaft JE, Rekhtman N, Lau C, Reva BA, Zakowski MF, Kris MG and Ladanyi M. EGFR exon 20 insertion mutations in lung adenocarcinomas: Prevalence, molecular heterogeneity, and clinicopathologic characteristics. Mol Cancer Ther. 2013; 12:220-229.

Balak MN, Gong Y, Riely GJ, Somwar R, Li AR, Zakowski MF, Chiang A, Yang G, Ouerfelli O, Kris MG, Ladanyi M, Miller VA and Pao W. Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor-mutant lung adenocarcinomas with acquired resistance to kinase inhibitors. Clin Cancer Res. 2006; 12:6494-6501.

Barquilla A and Pasquale EB. Eph receptors and ephrins: Therapeutic opportunities. Annu Rev Pharmacol Toxicol. 2015; 55:465-487.

Bean J, Brennan C, Shih JY, Riely G, Viale A, Wang L, Chitale D, Motoi N, Szoke J, Broderick S, Balak M, Chang WC, Yu CJ, Gazdar A, Pass H, Rusch V, Gerald W, Huang SF, Yang PC, Miller V, Ladanyi M, Yang CH and Pao W. MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. Proc Natl Acad Sci U S A. 2007; 104:20932-20937.

Belardinilli F, Capalbo C, Buffone A, Petroni M, Colicchia V, Ferraro S, Zani M, Nicolussi A, D'Inzeo S, Coppa A, Screpanti I, Gulino A and Giannini G. Validation of the ion torrent PGM sequencing for the prospective routine molecular diagnostic of colorectal cancer. Clin Biochem. 2015; 48:908-910.

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Chiara E Catenazzi M, Chang S, Neil Cooley R, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Huw Jones TA, Kang GD, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T, Ning Z, Ling Ng B, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos LL, Pickering L, Pike AC, Pike AC, Chris Pinkard D, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Ernest Sohna Sohna J, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Rogers J, Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R and Smith AJ. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008; 456:53-59.

Boch C, Kollmeier J, Roth A, Stephan-Falkenau S, Misch D, Gruning W, Bauer TT and Mairinger T. The frequency of EGFR and KRAS mutations in non-small cell lung cancer (NSCLC): Routine screening data for central Europe from a cohort study. BMJ Open. 2013; 3:10.1136/bmjopen-2013-002560. Print 2013.

Bosco N, Pelliccia F and Rocchi A. Characterization of FRA7B, a human common fragile site mapped at the 7p chromosome terminal region. Cancer Genet Cytogenet. 2010; 202:47-52.

Bott M, Brevet M, Taylor BS, Shimizu S, Ito T, Wang L, Creaney J, Lake RA, Zakowski MF, Reva B, Sander C, Delsite R, Powell S, Zhou Q, Shen R, Olshen A, Rusch V and Ladanyi M. The nuclear deubiquitinase BAP1 is commonly inactivated by somatic mutations and 3p21.1 losses in malignant pleural mesothelioma. Nat Genet. 2011; 43:668-672.

Brahmer J, Reckamp KL, Baas P, Crino L, Eberhardt WE, Poddubskaya E, Antonia S, Pluzanski A, Vokes EE, Holgado E, Waterhouse D, Ready N, Gainor J, Aren Frontera O, Havel L, Steins M, Garassino MC, Aerts JG, Domine M, Paz-Ares L, Reck M, Baudelet C, Harbison CT,

Lestini B and Spigel DR. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. N Engl J Med. 2015; 373:123-135.

Brannan JM, Dong W, Prudkin L, Behrens C, Lotan R, Bekele BN, Wistuba I and Johnson FM. Expression of the receptor tyrosine kinase EphA2 is increased in smokers and predicts poor survival in non-small cell lung cancer. Clin Cancer Res. 2009; 15:4423-4430.

Braslavsky I, Hebert B, Kartalov E and Quake SR. Sequence information can be obtained from single DNA molecules. Proc Natl Acad Sci U S A. 2003; 100:3960-3964.

Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, Roth R, George D, Eletr S, Albrecht G, Vermaas E, Williams SR, Moon K, Burcham T, Pallas M, DuBridge RB, Kirchner J, Fearon K, Mao J and Corcoran K. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat Biotechnol. 2000; 18:630-634.

Bria E, Pilotto S, Amato E, Fassan M, Novello S, Peretti U, Vavala T, Kinspergher S, Righi L, Santo A, Brunelli M, Corbo V, Giglioli E, Sperduti I, Milella M, Chilosi M, Scarpa A and Tortora G. Molecular heterogeneity assessment by next-generation sequencing and response to gefitinib of EGFR mutant advanced lung adenocarcinoma. Oncotarget. 2015; 6:12783-12795.

Brueckner B, Stresemann C, Kuner R, Mund C, Musch T, Meister M, Sultmann H and Lyko F. The human let-7a-3 locus contains an epigenetically regulated microRNA gene with oncogenic function. Cancer Res. 2007; 67:1419-1423.

Brustugun OT, Khattak AM, Tromborg AK, Beigi M, Beiske K, Lund-Iversen M and Helland A. BRAF-mutations in non-small cell lung cancer. Lung Cancer. 2014; 84:36-38.

Buikhuisen WA, Burgers JA, Vincent AD, Korse CM, van Klaveren RJ, Schramel FM, Pavlakis N, Nowak AK, Custers FL, Schouwink JH, Gans SJ, Groen HJ, Strankinga WF and Baas P. Thalidomide versus active supportive care for maintenance in patients with malignant mesothelioma after first-line chemotherapy (NVALT 5): An open-label, multicentre, randomised phase 3 study. Lancet Oncol. 2013; 14:543-551.

Cadby G, Mukherjee S, Musk AW, Reid A, Garlepp M, Dick I, Robinson C, Hui J, Fiorito G, Guarrera S, Beilby J, Melton PE, Moses EK, Ugolini D, Mirabelli D, Bonassi S, Magnani C, Dianzani I, Matullo G, Robinson B, Creaney J and Palmer LJ. A genome-wide association study for malignant mesothelioma risk. Lung Cancer. 2013; 82:1-8.

Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012; 489:519-525.

Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014; 513:202-209.

Cao X, Hong X, Jia X, Zhang L and Chen G. Single-nucleotide polymorphism rs41736 located in MET was significantly associated with prognosis of small cell lung cancer patients. Med Oncol. 2014; 31:333-014-0333-9. Epub 2014 Nov 22.

Carbone M, Ferris LK, Baumann F, Napolitano A, Lum CA, Flores EG, Gaudino G, Powers A, Bryant-Greenwood P, Krausz T, Hyjek E, Tate R, Friedberg J, Weigel T, Pass HI and Yang H. BAP1 cancer syndrome: Malignant mesothelioma, uveal and cutaneous melanoma, and MBAITs. J Transl Med. 2012; 10:179-5876-10-179.

Carrera S, Buque A, Azkona E, Aresti U, Calvo B, Sancho A, Arruti M, Nuno M, Rubio I, de Lobera AR, Lopez C and Vivanco GL. Epidermal growth factor receptor tyrosine-kinase inhibitor treatment resistance in non-small cell lung cancer: Biological basis and therapeutic strategies. Clin Transl Oncol. 2014; 16:339-350.

Caumont C, Veillon R, Gros A, Laharanne E, Begueret H and Merlio JP. Neuroendocrine phenotype as an acquired resistance mechanism in ALK-rearranged lung adenocarcinoma. Lung Cancer. 2016; 92:15-18.

Cha YJ, Cho BC, Kim HR, Lee HJ and Shim HS. A case of ALK-rearranged adenocarcinoma with small cell carcinoma-like transformation and resistance to crizotinib. J Thorac Oncol. 2016; 11:e55-8.

Chang TW. Binding of cells to matrixes of distinct antibodies coated on solid surface. J Immunol Methods. 1983; 65:217-223.

Chen XH, Shi H, Liu XL and Su B. The testis-specific apoptosis related gene TTL.6 underwent adaptive evolution in the lineage leading to humans. Gene. 2006; 370:58-63.

Chen D, Huang JF, Liu K, Zhang LQ, Yang Z, Chuai ZR, Wang YX, Shi DC, Huang Q and Fu WL. BRAFV600E mutation and its association with clinicopathological features of colorectal cancer: A systematic review and meta-analysis. PLoS One. 2014; 9:e90607.

Cheng NC, van Zandwijk N and Reid G. Cilengitide inhibits attachment and invasion of malignant pleural mesothelioma cells through antagonism of integrins alphavbeta3 and alphavbeta5. PLoS One. 2014; 9:e90374.

Chevrier S, Arnould L, Ghiringhelli F, Coudert B, Fumoleau P and Boidot R. Next-generation sequencing analysis of lung and colon carcinomas reveals a variety of genetic alterations. Int J Oncol. 2014; 45:1167-1174.

Choi Y, Sims GE, Murphy S, Miller JR and Chan AP. Predicting the functional effect of amino acid substitutions and indels. PLoS One. 2012; 7:e46688.

Chung JH, Rho JK, Xu X, Lee JS, Yoon HI, Lee CT, Choi YJ, Kim HR, Kim CH and Lee JC. Clinical and molecular evidences of epithelial to mesenchymal transition in acquired resistance to EGFR-TKIs. Lung Cancer. 2011; 73:176-182.

Ciavarella S, Milano A, Dammacco F and Silvestris F. Targeted therapies in cancer. BioDrugs. 2010; 24:77-88.

Cohen MH, Williams GA, Sridhara R, Chen G and Pazdur R. FDA drug approval summary: Gefitinib (ZD1839) (iressa) tablets. Oncologist. 2003; 8:303-306.

Cohen MH, Johnson JR, Chen YF, Sridhara R and Pazdur R. FDA drug approval summary: Erlotinib (tarceva) tablets. Oncologist. 2005; 10:461-466.

Coons AH, Creech HJ and Jones RN. Immunological properties of an antibody containing a fluorescent group. Proc Soc Exp Biol Med. 1941; 47:200-202.

Dahlman KB, Xia J, Hutchinson K, Ng C, Hucks D, Jia P, Atefi M, Su Z, Branch S, Lyle PL, Hicks DJ, Bozon V, Glaspy JA, Rosen N, Solit DB, Netterville JL, Vnencak-Jones CL, Sosman JA, Ribas A, Zhao Z and Pao W. BRAF(L597) mutations in melanoma are associated with sensitivity to MEK inhibitors. Cancer Discov. 2012; 2:791-797.

Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett MJ, Bottomley W, Davis N, Dicks E, Ewing R, Floyd Y, Gray K, Hall S, Hawes R, Hughes J, Kosmidou V, Menzies A, Mould C, Parker A, Stevens C, Watt S, Hooper S, Wilson R, Jayatilake H, Gusterson BA, Cooper C, Shipley J, Hargrave D, Pritchard-Jones K, Maitland N, Chenevix-Trench G, Riggins GJ, Bigner DD, Palmieri G, Cossu A, Flanagan A, Nicholson A, Ho JW, Leung SY, Yuen ST, Weber BL, Seigler HF, Darrow TL, Paterson H, Marais R, Marshall CJ, Wooster R, Stratton MR and Futreal PA. Mutations of the BRAF gene in human cancer. Nature. 2002; 417:949-954.

Davies H, Hunter C, Smith R, Stephens P, Greenman C, Bignell G, Teague J, Butler A, Edkins S, Stevens C, Parker A, O'Meara S, Avis T, Barthorpe S, Brackenbury L, Buck G, Clements J, Cole J, Dicks E, Edwards K, Forbes S, Gorton M, Gray K, Halliday K, Harrison R, Hills K, Hinton J, Jones D, Kosmidou V, Laman R, Lugg R, Menzies A, Perry J, Petty R, Raine K, Shepherd R, Small A, Solomon H, Stephens Y, Tofts C, Varian J, Webb A, West S, Widaa S, Yates A, Brasseur F, Cooper CS, Flanagan AM, Green A, Knowles M, Leung SY, Looijenga LH, Malkowicz B, Pierotti MA, Teh BT, Yuen ST, Lakhani SR, Easton DF, Weber BL, Goldstraw P, Nicholson AG, Wooster R, Stratton MR and Futreal PA. Somatic mutations of the protein kinase gene family in human lung cancer. Cancer Res. 2005; 65:7591-7595.

de Assis LV, Locatelli J and Isoldi MC. The role of key genes and pathways involved in the tumorigenesis of malignant mesothelioma. Biochim Biophys Acta. 2014a; 1845:232-247.

de Assis LV and Isoldi MC. The function, mechanisms, and role of the genes PTEN and TP53 and the effects of asbestos in the development of malignant mesothelioma: A review focused on the genes' molecular mechanisms. Tumour Biol. 2014b; 35:889-901.

de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, Yates L, Jamal-Hanjani M, Shafi S, Murugaesu N, Rowan AJ, Gronroos E, Muhammad MA, Horswell S, Gerlinger M, Varela I, Jones D, Marshall J, Voet T, Van Loo P, Rassl DM, Rintoul RC, Janes SM, Lee SM, Forster M, Ahmad T, Lawrence D, Falzon M, Capitanio A, Harkins TT, Lee CC, Tom W, Teefe E, Chen SC, Begum S, Rabinowitz A, Phillimore B, Spencer-Dene B, Stamp G, Szallasi Z, Matthews N, Stewart A, Campbell P and Swanton C. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. Science. 2014; 346:251-256.

De Falco V, Giannini R, Tamburrino A, Ugolini C, Lupi C, Puxeddu E, Santoro M and Basolo F. Functional characterization of the novel T599I-VKSRdel BRAF mutation in a follicular variant papillary thyroid carcinoma. J Clin Endocrinol Metab. 2008; 93:4398-4402.

de Groot P and Munden RF. Lung cancer epidemiology, risk factors, and prevention. Radiol Clin North Am. 2012; 50:863-876.

Dearden S, Stevens J, Wu YL and Blowers D. Mutation incidence and coincidence in non small-cell lung cancer: Meta-analyses by ethnicity and histology (mutMap). Ann Oncol. 2013; 24:2371-2376.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D and Daly MJ. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011; 43:491-498.

Devarakonda S, Morgensztern D and Govindan R. Genomic alterations in lung adenocarcinoma. Lancet Oncol. 2015; 16:e342-51.

Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, Sougnez C, Greulich H, Muzny DM, Morgan MB, Fulton L, Fulton RS, Zhang Q, Wendl MC, Lawrence MS, Larson DE, Chen K, Dooling DJ, Sabo A, Hawes AC, Shen H, Jhangiani SN, Lewis LR, Hall O, Zhu Y, Mathew T, Ren Y, Yao J, Scherer SE, Clerc K, Metcalf GA, Ng B, Milosavljevic A, Gonzalez-Garay ML, Osborne JR, Meyer R, Shi X, Tang Y, Koboldt DC, Lin L, Abbott R, Miner TL, Pohl C, Fewell G, Haipek C, Schmidt H, Dunford-Shore BH, Kraja A, Crosby SD, Sawyer CS, Vickery T, Sander S, Robinson J, Winckler W, Baldwin J, Chirieac LR, Dutt A, Fennell T, Hanna M, Johnson BE, Onofrio RC, Thomas RK, Tonon G, Weir BA, Zhao X, Ziaugra L, Zody MC, Giordano T, Orringer MB, Roth JA, Spitz MR, Wistuba II, Ozenberger B, Good PJ, Chang AC, Beer DG, Watson MA, Ladanyi M, Broderick S, Yoshizawa A, Travis

WD, Pao W, Province MA, Weinstock GM, Varmus HE, Gabriel SB, Lander ES, Gibbs RA, Meyerson M and Wilson RK. Somatic mutations affect key pathways in lung adenocarcinoma. Nature. 2008; 455:1069-1075.

Doebele RC, Pilling AB, Aisner DL, Kutateladze TG, Le AT, Weickhardt AJ, Kondo KL, Linderman DJ, Heasley LE, Franklin WA, Varella-Garcia M and Camidge DR. Mechanisms of resistance to crizotinib in patients with ALK gene rearranged non-small cell lung cancer. Clin Cancer Res. 2012; 18:1472-1482.

Dogan S, Shen R, Ang DC, Johnson ML, D'Angelo SP, Paik PK, Brzostowski EB, Riely GJ, Kris MG, Zakowski MF and Ladanyi M. Molecular epidemiology of EGFR and KRAS mutations in 3,026 lung adenocarcinomas: Higher susceptibility of women to smoking-related KRAS-mutant cancers. Clin Cancer Res. 2012; 18:6169-6177.

Dong X, Fernandez-Salas E, Li E and Wang S. Elucidation of resistance mechanisms to second-generation ALK inhibitors alectinib and ceritinib in non-small cell lung cancer cells. Neoplasia. 2016; 18:162-171.

Durkin SG and Glover TW. Chromosome fragile sites. Annu Rev Genet. 2007; 41:169-192.

Eletr ZM and Wilkinson KD. An emerging model for BAP1's role in regulating cell cycle progression. Cell Biochem Biophys. 2011; 60:3-11.

Elson JL, Smith PM, Greaves LC, Lightowlers RN, Chrzanowska-Lightowlers ZM, Taylor RW and Vila-Sanjurjo A. The presence of highly disruptive 16S rRNA mutations in clinical samples indicates a wider role for mutations of the mitochondrial ribosome in human disease. Mitochondrion. 2015; 25:17-27.

Engelman JA, Zejnullahu K, Mitsudomi T, Song Y, Hyland C, Park JO, Lindeman N, Gale CM, Zhao X, Christensen J, Kosaka T, Holmes AJ, Rogers AM, Cappuzzo F, Mok T, Lee C, Johnson BE, Cantley LC and Janne PA. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. Science. 2007; 316:1039-1043.

Engholm G, Ferlay J, Christensen N, Kejs AMT, Johannesen, TB, Khan S, Leinonen M, et al. NORDCAN: Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries http://www.ancr.nu . Accessed 10/15. Version 7.1 (2015).

Ettinger DS, Akerley W, Bepler G, Blum MG, Chang A, Cheney RT, Chirieac LR, D'Amico TA, Demmy TL, Ganti AK, Govindan R, Grannis FW,Jr, Jahan T, Jahanzeb M, Johnson DH, Kessinger A, Komaki R, Kong FM, Kris MG, Krug LM, Le QT, Lennes IT, Martins R, O'Malley J, Osarogiagbon RU, Otterson GA, Patel JD, Pisters KM, Reckamp K, Riely GJ, Rohren E, Simon GR, Swanson SJ, Wood DE, Yang SC and NCCN Non-Small Cell Lung Cancer Panel Members. Non-small cell lung cancer. J Natl Compr Canc Netw. 2010; 8:740-801.

Faoro L, Singleton PA, Cervantes GM, Lennon FE, Choong NW, Kanteti R, Ferguson BD, Husain AN, Tretiakova MS, Ramnath N, Vokes EE and Salgia R. EphA2 mutation in lung squamous cell carcinoma promotes increased cell survival, cell invasion, focal adhesions, and mammalian target of rapamycin activation. J Biol Chem. 2010; 285:18575-18585.

Farkona S, Diamandis EP and Blasutig IM. Cancer immunotherapy: the beginning of the end of cancer? BMC Med. 2016; 14:73.

Feng Z, Hu W, Hu Y and Tang MS. Acrolein is a major cigarette-related lung cancer agent: Preferential binding at p53 mutational hotspots and inhibition of DNA repair. Proc Natl Acad Sci U S A. 2006; 103:15404-15409.

Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D and Bray F. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015; 136:E359-86.

Fousteri M, Vermeulen W, van Zeeland AA and Mullenders LH. Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. Mol Cell. 2006; 23:471-482.

Frederiks CN, Lam SW, Guchelaar HJ and Boven E. Genetic polymorphisms and paclitaxel- or docetaxel-induced toxicities: A systematic review. Cancer Treat Rev. 2015; 41:935-950.

Fujita S, Masago K, Takeshita J, Okuda C, Otsuka K, Hata A, Kaji R, Katakami N and Hirata Y. Validation of an ion torrent sequencing platform for the detection of gene mutations in biopsy specimens from patients with non-small-cell lung cancer. PLoS One. 2015; 10:e0130219.

Fujita S, Masago K, Katakami N and Yatabe Y. Transformation to SCLC after treatment with the ALK inhibitor alectinib. J Thorac Oncol. 2016; [in press].

Gahr S, Stoehr R, Geissinger E, Ficker JH, Brueckl WM, Gschwendtner A, Gattenloehner S, Fuchs FS, Schulz C, Rieker RJ, Hartmann A, Ruemmele P and Dietmaier W. EGFR mutational status in a large series of caucasian european NSCLC patients: Data from daily practice. Br J Cancer. 2013; 109:1821-1828.

Gainor JF, Varghese AM, Ou SH, Kabraji S, Awad MM, Katayama R, Pawlak A, Mino-Kenudson M, Yeap BY, Riely GJ, Iafrate AJ, Arcila ME, Ladanyi M, Engelman JA, Dias-Santagata D and Shaw AT. ALK rearrangements are mutually exclusive with mutations in EGFR or KRAS: An analysis of 1,683 patients with non-small cell lung cancer. Clin Cancer Res. 2013; 19:4273-4281.

Galateau-Salle F, Churg A, Roggli V, Travis WD and World Health Organization Committee for Tumors of the Pleura. The 2015 world health organization classification of tumors of the pleura: Advances since the 2004 classification. J Thorac Oncol. 2016; 11:142-154.

Gerlinger M, McGranahan N, Dewhurst SM, Burrell RA, Tomlinson I and Swanton C. Cancer: Evolution within a lifetime. Annu Rev Genet. 2014; 48:215-236.

Giaginis C, Tsoukalas N, Bournakis E, Alexandrou P, Kavantzas N, Patsouris E and Theocharis S. Ephrin (eph) receptor A1, A4, A5 and A7 expression in human non-small cell lung carcinoma: Associations with clinicopathological parameters, tumor proliferative capacity and patients' survival. BMC Clin Pathol. 2014; 14:8-6890-14-8.

Goldberg AD, Allis CD and Bernstein E. Epigenetics: A landscape takes shape. Cell. 2007; 128:635-638.

Goparaju C, Donington JS, Hsu T, Harrington R, Hirsch N and Pass HI. Overexpression of EPH receptor B2 in malignant mesothelioma correlates with oncogenic behavior. J Thorac Oncol. 2013; 8:1203-1211.

Guan P, Yin Z, Li X, Wu W and Zhou B. Meta-analysis of human lung cancer microRNA expression profiling studies comparing cancer tissues with normal tissues. J Exp Clin Cancer Res. 2012; 31:54-9966-31-54.

Guo G, Chmielecki J, Goparaju C, Heguy A, Dolgalev I, Carbone M, Seepo S, Meyerson M and Pass HI. Whole-exome sequencing reveals frequent genetic alterations in BAP1, NF2, CDKN2A, and CUL1 in malignant pleural mesothelioma. Cancer Res. 2015; 75:264-269.

Hadd AG, Houghton J, Choudhary A, Sah S, Chen L, Marko AC, Sanford T, Buddavarapu K, Krosting J, Garmire L, Wylie D, Shinde R, Beaudenon S, Alexander EK, Mambo E, Adai AT and Latham GJ. Targeted, high-depth, next-generation sequencing of cancer genes in formalin-

fixed, paraffin-embedded and fine-needle aspiration tumor specimens. J Mol Diagn. 2013; 15:234-247.

Hahn Y, Lee YJ, Yun JH, Yang SK, Park CW, Mita K, Huh TL, Rhee M and Chung JH. Duplication of genes encoding non-clathrin coat protein gamma-COP in vertebrate, insect and plant evolution. FEBS Lett. 2000; 482:31-36.

Han JY, Kim SH, Lee YS, Lee SY, Hwang JA, Kim JY, Yoon SJ and Lee GK. Comparison of targeted next-generation sequencing with conventional sequencing for predicting the responsiveness to epidermal growth factor receptor-tyrosine kinase inhibitor (EGFR-TKI) therapy in never-smokers with lung adenocarcinoma. Lung Cancer. 2014; 85:161-167.

Hanahan D and Weinberg RA. Hallmarks of cancer: The next generation. Cell. 2011; 144:646-674.

Hancock DG, Langley ME, Chia KL, Woodman RJ and Shanahan EM. Wood dust exposure and lung cancer risk: A meta-analysis. Occup Environ Med. 2015; 72:889-898.

Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, Causey M, Colonell J, Dimeo J, Efcavitch JW, Giladi E, Gill J, Healy J, Jarosz M, Lapen D, Moulton K, Quake SR, Steinmann K, Thayer E, Tyurina A, Ward R, Weiss H and Xie Z. Single-molecule DNA sequencing of a viral genome. Science. 2008; 320:106-109.

Hashim D, Boffetta P, La Vecchia C, Rota M, Bertuccio P, Malvezzi M and Negri E. The global decrease in cancer mortality: Trends and disparities. Ann Oncol. 2016; 27:926-933.

Hata A, Yoshioka H, Fujita S, Kunimasa K, Kaji R, Imai Y, Tomii K, Iwasaku M, Nishiyama A, Ishida T and Katakami N. Complex mutations in the epidermal growth factor receptor gene in non-small cell lung cancer. J Thorac Oncol. 2010; 5:1524-1528.

Heather JM and Chain B. The sequence of sequencers: The history of sequencing DNA. Genomics. 2016; 107:1-8.

Hiddinga BI, Rolfo C and van Meerbeeck JP. Mesothelioma treatment: Are we on target? A review. J Adv Res. 2015; 6:319-330.

Hirota E, Yan L, Tsunoda T, Ashida S, Fujime M, Shuin T, Miki T, Nakamura Y and Katagiri T. Genome-wide gene expression profiles of clear cell renal cell carcinoma: Identification of molecular targets for treatment of renal cell carcinoma. Int J Oncol. 2006; 29:799-827.

Hollstein M, Sidransky D, Vogelstein B and Harris CC. P53 mutations in human cancers. Science. 1991; 253:49-53.

Hsieh MH, Fang YF, Chang WC, Kuo HP, Lin SY, Liu HP, Liu CL, Chen HC, Ku YC, Chen YT, Chang YH, Chen YT, Hsi BL, Tsai SF and Huang SF. Complex mutation patterns of epidermal growth factor receptor gene associated with variable responses to gefitinib treatment in patients with non-small cell lung cancer. Lung Cancer. 2006; 53:311-322.

Husa AM, MagiA AE, Larsson M, Fornander T and Perez-Tenorio G. EPH/ephrin profile and EPHB2 expression predicts patient survival in breast cancer. Oncotarget. 2016; [in press].

Husgafvel-Pursiainen K, Karjalainen A, Kannio A, Anttila S, Partanen T, Ojajarvi A and Vainio H. Lung cancer and past occupational exposure to asbestos. Role of p53 and K-ras mutations. Am J Respir Cell Mol Biol. 1999; 20:667-674.

IARC. IARC monographs on the evaluation of carcinogenic risks to humans. Personal habits and indoor combustions: Tobacco smoking. 2012a; 100E:43-211.

IARC. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. A Review of Human Carcinogens. Arsenic, Metals, Fibres and Dusts. 2012b; 100C:219-309.

Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, Cho J, Suh J, Capelletti M, Sivachenko A, Sougnez C, Auclair D, Lawrence MS, Stojanov P, Cibulskis K, Choi K, de

Waal L, Sharifnia T, Brooks A, Greulich H, Banerji S, Zander T, Seidel D, Leenders F, Ansen S, Ludwig C, Engel-Riedel W, Stoelben E, Wolf J, Goparju C, Thompson K, Winckler W, Kwiatkowski D, Johnson BE, Janne PA, Miller VA, Pao W, Travis WD, Pass HI, Gabriel SB, Lander ES, Thomas RK, Garraway LA, Getz G and Meyerson M. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell. 2012; 150:1107-1120.

International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature. 2004; 431:931-945.

Isozaki H, Ichihara E, Takigawa N, Ohashi K, Ochi N, Yasugi M, Ninomiya T, Yamane H, Hotta K, Sakai K, Matsumoto K, Hosokawa S, Bessho A, Sendo T, Tanimoto M and Kiura K. Non-small cell lung cancer cells acquire resistance to the ALK inhibitor alectinib by activating alternative receptor tyrosine kinases. Cancer Res. 2016; 76:1506-1516.

Ivetac I, Gurung R, Hakim S, Horan KA, Sheffield DA, Binge LC, Majerus PW, Tiganis T and Mitchell CA. Regulation of PI(3)K/Akt signalling and cellular transformation by inositol polyphosphate 4-phosphatase-1. EMBO Rep. 2009; 10:487-493.

Jamal S, Cheriyan VT, Muthu M, Munie S, Levi E, Ashour AE, Pass HI, Wali A, Singh M, Rishi AK. CARP-1 functional mimetics are novel class of small molecule inhibitors of malignant pleural mesothelioma cells. PLoS One. 2014; 9:e89146.

Janne PA, Yang JC, Kim DW, Planchard D, Ohe Y, Ramalingam SS, Ahn MJ, Kim SW, Su WC, Horn L, Haggstrom D, Felip E, Kim JH, Frewer P, Cantarini M, Brown KH, Dickinson PA, Ghiorghiu S and Ranson M. AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. N Engl J Med. 2015; 372:1689-1699.

Ji XD, Li G, Feng YX, Zhao JS, Li JJ, Sun ZJ, Shi S, Deng YZ, Xu JF, Zhu YQ, Koeffler HP, Tong XJ and Xie D. EphB3 is overexpressed in non-small-cell lung cancer and promotes tumor metastasis by enhancing cell survival and migration. Cancer Res. 2011; 71:1156-1166.

Johnson DH, Fehrenbacher L, Novotny WF, Herbst RS, Nemunaitis JJ, Jablons DM, Langer CJ, DeVore RF,3rd, Gaudreault J, Damico LA, Holmgren E and Kabbinavar F. Randomized phase II trial comparing bevacizumab plus carboplatin and paclitaxel with carboplatin and paclitaxel alone in previously untreated locally advanced or metastatic non-small-cell lung cancer. J Clin Oncol. 2004; 22:2184-2191.

Juhlin CC, Goh G, Healy JM, Fonseca AL, Scholl UI, Stenman A, Kunstman JW, Brown TC, Overton JD, Mane SM, Nelson-Williams C, Backdahl M, Suttorp AC, Haase M, Choi M, Schlessinger J, Rimm DL, Hoog A, Prasad ML, Korah R, Larsson C, Lifton RP and Carling T. Whole-exome sequencing characterizes the landscape of somatic mutations and copy number alterations in adrenocortical carcinoma. J Clin Endocrinol Metab. 2015; 100:E493-502.

Kadara H, Scheet P, Wistuba II and Spira AE. Early events in the molecular pathogenesis of lung cancer. Cancer Prev Res (Phila). 2016; [in press].

Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F and Pinkel D. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science. 1992; 258:818-821.

Kang HC, Kim HK, Lee S, Mendez P, Kim JW, Woodard G, Yoon JH, Jen KY, Fang LT, Jones K, Jablons DM and Kim IJ. Whole exome and targeted deep sequencing identify genome-wide allelic loss and frequent SETDB1 mutations in malignant pleural mesotheliomas. Oncotarget. 2016; 7:8321-8331.

Karam-Hage M, Cinciripini PM and Gritz ER. Tobacco use and cessation for cancer survivors: An overview for clinicians. CA Cancer J Clin. 2014; 64:272-290.

Katayama R, Shaw AT, Khan TM, Mino-Kenudson M, Solomon BJ, Halmos B, Jessop NA, Wain JC, Yeo AT, Benes C, Drew L, Saeh JC, Crosby K, Sequist LV, Iafrate AJ and Engelman JA. Mechanisms of acquired crizotinib resistance in ALK-rearranged lung cancers. Sci Transl Med. 2012; 4:120ra17.

Katayama R, Friboulet L, Koike S, Lockerman EL, Khan TM, Gainor JF, Iafrate AJ, Takeuchi K, Taiji M, Okuno Y, Fujita N, Engelman JA and Shaw AT. Two novel ALK mutations mediate acquired resistance to the next-generation ALK inhibitor alectinib. Clin Cancer Res. 2014; 20:5686-5696.

Kazandjian D, Suzman DL, Blumenthal G, Mushti S, He K, Libeg M, Keegan P and Pazdur R. FDA approval summary: Nivolumab for the treatment of metastatic non-small cell lung cancer with progression on or after platinum-based chemotherapy. Oncologist. 2016; 21:634-642.

Keam B, Kim DW, Park JH, Lee JO, Kim TM, Lee SH, Chung DH and Heo DS. Rare and complex mutations of epidermal growth factor receptor, and efficacy of tyrosine kinase inhibitor in patients with non-small cell lung cancer. Int J Clin Oncol. 2014; 19:594-600.

Keir ME, Butte MJ, Freeman GJ and Sharpe AH. PD-1 and its ligands in tolerance and immunity. Annu Rev Immunol. 2008; 26:677-704.

Kettunen E, Knuutila S. Occupational Cancers. Malignant mesothelioma: molecular markers. Springer-Verlag, London. 2014: 325-343.

Khodayari N, Mohammed KA, Goldberg EP and Nasreen N. EphrinA1 inhibits malignant mesothelioma tumor growth via let-7 microRNA-mediated repression of the RAS oncogene. Cancer Gene Ther. 2011; 18:806-816.

Khozin S, Blumenthal GM, Zhang L, Tang S, Brower M, Fox E, Helms W, Leong R, Song P, Pan Y, Liu Q, Zhao P, Zhao H, Lu D, Tang Z, Al Hakim A, Boyd K, Keegan P, Justice R and Pazdur R. FDA approval: Ceritinib for the treatment of metastatic anaplastic lymphoma kinase-positive non-small cell lung cancer. Clin Cancer Res. 2015; 21:2436-2439.

Kim JH, Chu SC, Gramlich JL, Pride YB, Babendreier E, Chauhan D, Salgia R, Podar K, Griffin JD and Sattler M. Activation of the PI3K/mTOR pathway by BCR-ABL contributes to increased production of reactive oxygen species. Blood. 2005; 105:1717-1723.

Kim EY, Cho EN, Park HS, Hong JY, Lim S, Youn JP, Hwang SY and Chang YS. Compound EGFR mutation is frequently detected with co-mutations of actionable genes and associated with poor clinical outcome in lung adenocarcinoma. Cancer Biol Ther. 2016; 17:237-245.

Knudson AG,Jr. Mutation and cancer: Statistical study of retinoblastoma. Proc Natl Acad Sci U S A. 1971; 68:820-823.

Kobayashi S, Canepa HM, Bailey AS, Nakayama S, Yamaguchi N, Goldstein MA, Huberman MS and Costa DB. Compound EGFR mutations and response to EGFR tyrosine kinase inhibitors. J Thorac Oncol. 2013; 8:45-51.

Koivunen JP, Kim J, Lee J, Rogers AM, Park JO, Zhao X, Naoki K, Okamoto I, Nakagawa K, Yeap BY, Meyerson M, Wong KK, Richards WG, Sugarbaker DJ, Johnson BE and Janne PA. Mutations in the LKB1 tumour suppressor are frequently detected in tumours from Caucasian but not Asian lung cancer patients. Br J Cancer. 2008; 99:245-252.

Krug LM, Kindler HL, Calvert H, Manegold C, Tsao AS, Fennell D, Ohman R, Plummer R, Eberhardt WE, Fukuoka K, Gaafar RM, Lafitte JJ, Hillerdal G, Chu Q, Buikhuisen WA, Lubiniecki GM, Sun X, Smith M and Baas P. Vorinostat in patients with advanced malignant pleural mesothelioma who have progressed on previous chemotherapy (VANTAGE-014): A phase 3, double-blind, randomised, placebo-controlled trial. Lancet Oncol. 2015; 16:447-456.

Kuan FC, Kuo LT, Chen MC, Yang CT, Shi CS, Teng D and Lee KD. Overall survival benefits of first-line EGFR tyrosine kinase inhibitors in EGFR-mutated non-small-cell lung cancers: A systematic review and meta-analysis. Br J Cancer. 2015; 113:1519-1528.

Kumar P, Henikoff S and Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009; 4:1073-1081.

Lahtela J, Corson LB, Hemmes A, Brauer MJ, Koopal S, Lee J, Hunsaker TL, Jackson PK and Verschuren EW. A high-content cellular senescence screen identifies candidate tumor suppressors, including EPHA3. Cell Cycle. 2013; 12:625-634.

Lahtela J, Pradhan B, Narhi K, Hemmes A, Sarkioja M, Kovanen PE, Brown A and Verschuren EW. The putative tumor suppressor gene EphA3 fails to demonstrate a crucial role in murine lung tumorigenesis or morphogenesis. Dis Model Mech. 2015; 8:393-401.

Laitinen VH, Akinrinade O, Rantapero T, Tammela TL, Wahlfors T and Schleutker J. Germline copy number variation analysis in Finnish families with hereditary prostate cancer. Prostate. 2016; 76:316-324.

Langer-Safer PR, Levine M and Ward DC. Immunological method for mapping genes on drosophila polytene chromosomes. Proc Natl Acad Sci U S A. 1982; 79:4381-4385.

Langevin SM, Kratzke RA and Kelsey KT. Epigenetics of lung cancer. Transl Res. 2015; 165:74-90.

Langmead B and Salzberg SL. Fast gapped-read alignment with bowtie 2. Nat Methods. 2012; 9:357-359.

Lee CK, Wu YL, Ding PN, Lord SJ, Inoue A, Zhou C, Mitsudomi T, Rosell R, Pavlakis N, Links M, Gebski V, Gralla RJ and Yang JC. Impact of specific epidermal growth factor receptor (EGFR) mutations and clinical characteristics on outcomes after treatment with EGFR tyrosine kinase inhibitors versus chemotherapy in EGFR-mutant lung cancer: A meta-analysis. J Clin Oncol. 2015; 33:1958-1965.

Lee T, Lee B, Choi Y, Han J, Ahn MJ and Um SW. Non-small cell lung cancer with concomitant EGFR, KRAS, and ALK mutation: Clinicopathologic features of 12 cases. J Pathol Transl Med. 2016; [in press].

Lemjabbar-Alaoui H, Hassan OU, Yang YW and Buchanan P. Lung cancer: Biology and treatment options. Biochim Biophys Acta. 2015; 1856:189-210.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R and 1000 Genome Project Data Processing Subgroup. The sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078-2079.

Li H and Durbin R. Fast and accurate long-read alignment with burrows-wheeler transform. Bioinformatics. 2010; 26:589-595.

Li H, Pan Y, Li Y, Li C, Wang R, Hu H, Zhang Y, Ye T, Wang L, Shen L, Sun Y and Chen H. Frequency of well-identified oncogenic driver mutations in lung adenocarcinoma of smokers varies with histological subtypes and graduated smoking dose. Lung Cancer. 2013; 79:8-13.

Li RX, Chen ZH and Chen ZK. The role of EPH receptors in cancer-related epithelial-mesenchymal transition. Chin J Cancer. 2014; 33:231-240.

Li S, Li L, Zhu Y, Huang C, Qin Y, Liu H, Ren-Heidenreich L, Shi B, Ren H, Chu X, Kang J, Wang W, Xu J, Tang K, Yang H, Zheng Y, He J, Yu G and Liang N. Coexistence of EGFR with KRAS, or BRAF, or PIK3CA somatic mutations in lung cancer: A comprehensive mutation profiling from 5125 chinese cohorts. Br J Cancer. 2014; 110:2812-2820.

Li L, Chen BF and Chan WY. An epigenetic regulator: Methyl-CpG-binding domain protein 1 (MBD1). Int J Mol Sci. 2015; 16:5125-5140.

Lim JU, Yeo CD, Rhee CK, Kim YH, Park CK, Kim JS, Kim JW, Lee SH, Kim SJ, Yoon HK, Kim TJ and Lee KY. Chronic obstructive pulmonary disease-related non-small-cell lung cancer exhibits a low prevalence of EGFR and ALK driver mutations. PLoS One. 2015; 10:e0142306.

Lin MT, Mosier SL, Thiess M, Beierl KF, Debeljak M, Tseng LH, Chen G, Yegnasubramanian S, Ho H, Cope L, Wheelan SJ, Gocke CD and Eshleman JR. Clinical validation of KRAS, BRAF, and EGFR mutation detection using next-generation sequencing. Am J Clin Pathol. 2014; 141:856-866.

Lisabeth EM, Falivelli G and Pasquale EB. Eph receptor signaling and ephrins. Cold Spring Harb Perspect Biol. 2013; 5:10.1101/cshperspect.a009159.

Lisabeth EM, Fernandez C and Pasquale EB. Cancer somatic mutations disrupt functions of the EphA3 receptor tyrosine kinase through multiple mechanisms. Biochemistry. 2012; 51:1464-1475.

Liu H, Jin G, Wang H, Wu W, Liu Y, Qian J, Fan W, Ma H, Miao R, Hu Z, Sun W, Wang Y, Jin L, Wei Q, Shen H, Huang W and Lu D. Methyl-CpG binding domain 1 gene polymorphisms and lung cancer risk in a chinese population. Biomarkers. 2008; 13:607-617.

Liu R, Ferguson BD, Zhou Y, Naga K, Salgia R, Gill PS and Krasnoperov V. EphB4 as a therapeutic target in mesothelioma. BMC Cancer. 2013; 13:269-2407-13-269.

Liu X, Lin XJ, Wang CP, Yan KK, Zhao LY, An WX and Liu XD. Association between smoking and p53 mutation in lung cancer: A meta-analysis. Clin Oncol (R Coll Radiol). 2014; 26:18-24.

Liu S, Meric-Bernstam F, Parinyanitikul N, Wang B, Eterovic AK, Zheng X, Gagea M, Chavez-MacGregor M, Ueno NT, Lei X, Zhou W, Nair L, Tripathy D, Brown PH, Hortobagyi GN, Chen K, Mendelsohn J, Mills GB and Gonzalez-Angulo AM. Functional consequence of the MET-T1010I polymorphism in breast cancer. Oncotarget. 2015; 6:2604-2614.

Liu X, Xu Y, Jin Q, Wang W, Zhang S, Wang X, Zhang Y, Xu X and Huang J. EphA8 is a prognostic marker for epithelial ovarian cancer. Oncotarget. 2016a; [in press].

Liu J, Sheng Z, Zhang Y and Li G. The efficacy of epidermal growth factor receptor tyrosine kinase inhibitors for molecularly selected patients with non-small cell lung cancer: A meta-analysis of 30 randomized controlled trials. Target Oncol. 2016b; 11:49-58.

Lo Iacono M, Monica V, Righi L, Grosso F, Libener R, Vatrano S, Bironzo P, Novello S, Musmeci L, Volante M, Papotti M and Scagliotti GV. Targeted next-generation sequencing of cancer genes in advanced stage malignant pleural mesothelioma: A retrospective study. J Thorac Oncol. 2015; 10:492-499.

Lynch TJ, Bondarenko I, Luft A, Serwatowski P, Barlesi F, Chacko R, Sebastian M, Neal J, Lu H, Cuillerot JM and Reck M. Ipilimumab in combination with paclitaxel and carboplatin as first-line treatment in stage IIIB/IV non-small-cell lung cancer: Results from a randomized, double-blind, multicenter phase II study. J Clin Oncol. 2012; 30:2046-2054.

Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, Louis DN, Christiani DC, Settleman J and Haber DA. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med. 2004; 350:2129-2139.

Ma PC, Jagadeeswaran R, Jagadeesh S, Tretiakova MS, Nallasura V, Fox EA, Hansen M, Schaefer E, Naoki K, Lader A, Richards W, Sugarbaker D, Husain AN, Christensen JG and Salgia R. Functional expression and mutations of c-met and its therapeutic inhibition with SU11274 and small interfering RNA in non-small cell lung cancer. Cancer Res. 2005; 65:1479-1488.

Maki-Nevala S, Knuuttila A, Knuutila S and Sarhadi VK. Concordant results of epidermal growth factor receptor mutation detection by real-time polymerase chain reaction and ion torrent technology in non-small cell lung cancer. J Clin Respir Dis Care. 2016; 2:107.

Malapelle U, Vigliar E, Sgariglia R, Bellevicine C, Colarossi L, Vitale D, Pallante P and Troncone G. Ion torrent next-generation sequencing for routine identification of clinically relevant mutations in colorectal cancer patients. J Clin Pathol. 2015; 68:64-68.

Malik SM, Maher VE, Bijwaard KE, Becker RL, Zhang L, Tang SW, Song P, Liu Q, Marathe A, Gehrke B, Helms W, Hanner D, Justice R and Pazdur R. U.S. food and drug administration approval: Crizotinib for treatment of advanced or metastatic non-small cell lung cancer that is anaplastic lymphoma kinase positive. Clin Cancer Res. 2014; 20:2029-2034.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF and Rothberg JM. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 2005; 437:376-380.

Mazieres J, Barlesi F, Filleron T, Besse B, Monnet I, Beau-Faller M, Peters S, Dansin E, Fruh M, Pless M, Rosell R, Wislez M, Fournel P, Westeel V, Cappuzzo F, Cortot A, Moro-Sibilot D, Milia J and Gautschi O. Lung cancer patients with HER2 mutations treated with chemotherapy and HER2-targeted drugs: Results from the european EUHER2 cohort. Ann Oncol. 2016; 27:281-286.

Meng F, Henson R, Wehbe-Janek H, Ghoshal K, Jacob ST and Patel T. MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer. Gastroenterology. 2007; 133:647-658.

Miao B, Ji Z, Tan L, Taylor M, Zhang J, Choi HG, Frederick DT, Kumar R, Wargo JA, Flaherty KT, Gray NS and Tsao H. EPHA2 is a mediator of vemurafenib resistance and a novel therapeutic target in melanoma. Cancer Discov. 2015; 5:274-287.

Midha A, Dearden S and McCormack R. EGFR mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: A systematic review and global map by ethnicity (mutMapII). Am J Cancer Res. 2015; 5:2892-2911.

Minguet J, Smith KH and Bramlage P. Targeted therapies for treatment of non-small cell lung cancer-recent advances and future perspectives. Int J Cancer. 2016; 138:2549-2561.

Mitsudomi T, Morita S, Yatabe Y, Negoro S, Okamoto I, Tsurutani J, Seto T, Satouchi M, Tada H, Hirashima T, Asami K, Katakami N, Takada M, Yoshioka H, Shibata K, Kudoh S, Shimizu E, Saito H, Toyooka S, Nakagawa K, Fukuoka M and West Japan Oncology Group. Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): An open label, randomised phase 3 trial. Lancet Oncol. 2010; 11:121-128.

Miura JT, Johnston FM, Gamblin TC and Turaga KK. Current trends in the management of malignant peritoneal mesothelioma. Ann Surg Oncol. 2014; 21:3947-3953.

Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, Sunpaweravong P, Han B, Margono B, Ichinose Y, Nishiwaki Y, Ohe Y, Yang JJ, Chewaskulyong B, Jiang H, Duffield EL, Watkins

CL, Armour AA and Fukuoka M. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009; 361:947-957.

Morera L, Lubbert M and Jung M. Targeting histone methyltransferases and demethylases in clinical trials for cancer therapy. Clin Epigenetics. 2016; 8:57.

Muller IB, De Langen AJ, Honeywell RJ, Giovannetti E and Peters GJ. Overcoming crizotinib resistance in ALK-rearranged NSCLC with the second-generation ALK-inhibitor ceritinib. Expert Rev Anticancer Ther. 2016; 16:147-157.

Mullis KB. Target amplification for DNA analysis by the polymerase chain reaction. Ann Biol Clin (Paris). 1990; 48:579-582.

Nasreen N, Mohammed KA, Lai Y and Antony VB. Receptor EphA2 activation with ephrinA1 suppresses growth of malignant mesothelioma (MM). Cancer Lett. 2007; 258:215-222.

Nasreen N, Mohammed KA and Antony VB. Silencing the receptor EphA2 suppresses the growth and haptotaxis of malignant mesothelioma cells. Cancer. 2006; 107:2425-2435.

Newton CR, Graham A, Heptinstall LE, Powell SJ, Summers C, Kalsheker N, Smith JC and Markham AF. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). Nucleic Acids Res. 1989; 17:2503-2516.

Nigorikawa K, Hazeki K, Sasaki J, Omori Y, Miyake M, Morioka S, Guo Y, Sasaki T and Hazeki O. Inositol polyphosphate-4-phosphatase type I negatively regulates phagocytosis via dephosphorylation of phagosomal PtdIns(3,4)P2. PLoS One. 2015; 10:e0142091.

Ou SH, Greenbowe J, Khan ZU, Azada MC, Ross JS, Stevens PJ, Ali SM, Miller VA and Gitlitz B. I1171 missense mutation (particularly I1171N) is a common resistance mutation in ALK-positive NSCLC patients who have progressive disease while on alectinib and is sensitive to ceritinib. Lung Cancer. 2015; 88:231-234.

Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, Naoki K, Sasaki H, Fujii Y, Eck MJ, Sellers WR, Johnson BE and Meyerson M. EGFR mutations in lung cancer: Correlation with clinical response to gefitinib therapy. Science. 2004; 304:1497-1500.

Pallis AG and Syrigos KN. Lung cancer in never smokers: Disease characteristics and risk factors. Crit Rev Oncol Hematol. 2013; 88:494-503.

Pan W, Yang Y, Zhu H, Zhang Y, Zhou R and Sun X. KRAS mutation is a weak, but valid predictor for poor prognosis and treatment outcomes in NSCLC: A meta-analysis of 41 studies. Oncotarget. 2016; 7:8373-8388.

Pao W, Miller VA, Politi KA, Riely GJ, Somwar R, Zakowski MF, Kris MG and Varmus H. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. PLoS Med. 2005; 2:e73.

Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, Singh B, Heelan R, Rusch V, Fulton L, Mardis E, Kupfer D, Wilson R, Kris M and Varmus H. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci U S A. 2004; 101:13306-13311.

Park HJ, Kim BG, Lee SJ, Heo SH, Kim JY, Kwon TH, Lee EB, Ryoo HM and Cho JY. Proteomic profiling of endothelial cells in human lung cancer. J Proteome Res. 2008; 7:1138-1150.

Pasquale EB. Eph receptors and ephrins in cancer: Bidirectional signalling and beyond. Nat Rev Cancer. 2010; 10:165-180.

Patel JN, Ersek JL and Kim ES. Lung cancer biomarkers, targeted therapies and clinical assays. Transl Lung Cancer Res. 2015; 4:503-514.

Pecuchet N, Laurent-Puig P, Mansuet-Lupo A, Legras A, Alifano M, Pallier K, Didelot A, Gibault L, Danel C, Just PA, Riquet M, Le Pimpec-Barthes F, Damotte D, Fabre E and Blons H. Different prognostic impact of STK11 mutations in non-squamous non-small-cell lung cancer. Oncotarget. 2015; [in press].

Pei N, Cao L, Liu Y, Wu J, Song Q, Zhang Z, Yuan J and Zhang X. XAB2 tagSNPs contribute to non-small cell lung cancer susceptibility in Chinese population. BMC Cancer. 2015; 15:560-015-1567-4.

Pekar-Zlotin M, Hirsch FR, Soussan-Gutman L, Ilouze M, Dvir A, Boyle T, Wynes M, Miller VA, Lipson D, Palmer GA, Ali SM, Dekel S, Brenner R, Bunn PA,Jr and Peled N. Fluorescence in situ hybridization, immunohistochemistry, and next-generation sequencing for detection of EML4-ALK rearrangement in lung cancer. Oncologist. 2015; 20:316-322.

Peng L, Song Z and Jiao S. Comparison of uncommon EGFR exon 21 L858R compound mutations with single mutation. Onco Targets Ther. 2015; 8:905-910.

Pillai RN and Ramalingam SS. Heat shock protein 90 inhibitors in non-small-cell lung cancer. Curr Opin Oncol. 2014; 26:159-164.

Pirker R. Epidermal growth factor receptor-directed monoclonal antibodies in nonsmall cell lung cancer: an update. Curr Opin Oncol. 2015; 27:87-93.

Pirker R. Third-generation epidermal growth factor receptor tyrosine kinase inhibitors in advanced nonsmall cell lung cancer. Curr Opin Oncol. 2016; 28:115-121.

Poulikakos PI, Xiao GH, Gallagher R, Jablonski S, Jhanwar SC and Testa JR. Re-expression of the tumor suppressor NF2/merlin inhibits invasiveness in mesothelioma cells and negatively regulates FAK. Oncogene. 2006; 25:5960-5968.

Pujol JL, Pirker R, Lynch TJ, Butts CA, Rosell R, Shepherd FA, Vansteenkiste J, O'Byrne KJ, de Blas B, Heighway J, von Heydebreck A and Thatcher N. Meta-analysis of individual patient data from randomized trials of chemotherapy plus cetuximab as first-line treatment for advanced non-small cell lung cancer. Lung Cancer. 2014; 83:211-218.

Pushkarev D, Neff NF and Quake SR. Single-molecule sequencing of an individual human genome. Nat Biotechnol. 2009; 27:847-850.

Ramlau R, Cufer T, Berzinec P, Dziadziuszko R, Olszewski W, Popper H, Bajcic P, Dusek L, Zbozinkova Z, Pirker R and INSIGHT study team. Epidermal growth factor receptor mutation-positive non-small-cell lung cancer in the real-world setting in central Europe: The INSIGHT study. J Thorac Oncol. 2015; 10:1370-1374.

Rauch TA, Zhong X, Wu X, Wang M, Kernstine KH, Wang Z, Riggs AD and Pfeifer GP. High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. Proc Natl Acad Sci U S A. 2008; 105:252-257.

Reck M, von Pawel J, Zatloukal P, Ramlau R, Gorbounova V, Hirsh V, Leighl N, Mezger J, Archer V, Moore N and Manegold C. Phase III trial of cisplatin plus gemcitabine with either placebo or bevacizumab as first-line therapy for nonsquamous non-small-cell lung cancer: AVAil. J Clin Oncol. 2009; 27:1227-1234.

Reck M, von Pawel J, Zatloukal P, Ramlau R, Gorbounova V, Hirsh V, Leighl N, Mezger J, Archer V, Moore N, Manegold C and BO17704 Study Group. Overall survival with cisplatin-gemcitabine and bevacizumab or placebo as first-line therapy for nonsquamous non-small-cell lung cancer: Results from a randomised phase III trial (AVAiL). Ann Oncol. 2010; 21:1804-1809.

Reck M, Kaiser R, Mellemgaard A, Douillard JY, Orlov S, Krzakowski M, von Pawel J, Gottfried M, Bondarenko I, Liao M, Gann CN, Barrueco J, Gaschler-Markefski B, Novello S and LUME-Lung 1 Study Group. Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-lung 1): A phase 3, double-blind, randomised controlled trial. Lancet Oncol. 2014; 15:143-155.

Rideout WM,3rd, Coetzee GA, Olumi AF and Jones PA. 5-methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. Science. 1990; 249:1288-1290.

Rivlin N, Brosh R, Oren M and Rotter V. Mutations in the p53 tumor suppressor gene: Important milestones at the various steps of tumorigenesis. Genes Cancer. 2011; 2:466-474.

Ronaghi M, Uhlen M and Nyren P. A sequencing method based on real-time pyrophosphate. Science. 1998; 281:363, 365.

Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, Hoon J, Simons JF, Marran D, Myers JW, Davidson JF, Branting A, Nobile JR, Puc BP, Light D, Clark TA, Huber M, Branciforte JT, Stoner IB, Cawley SE, Lyons M, Fu Y, Homer N, Sedova M, Miao X, Reed B, Sabina J, Feierstein E, Schorn M, Alanjary M, Dimalanta E, Dressman D, Kasinskas R, Sokolsky T, Fidanza JA, Namsaraev E, McKernan KJ, Williams A, Roth GT and Bustillo J. An integrated semiconductor device enabling non-optical genome sequencing. Nature. 2011; 475:348-352.

Russell PA, Barnett SA, Walkiewicz M, Wainer Z, Conron M, Wright GM, Gooi J, Knight S, Wynne R, Liew D and John T. Correlation of mutation status and survival with predominant histologic subtype according to the new IASLC/ATS/ERS lung adenocarcinoma classification in stage III (N2) patients. J Thorac Oncol. 2013; 8:461-468.

Saber A, van der Wekken AJ, Kerner GS, van den Berge M, Timens W, Schuuring E, Ter Elst A, van den Berg A, Hiltermann TJ and Groen HJ. Chronic obstructive pulmonary disease is not associated with KRAS mutations in non-small cell lung cancer. PLoS One. 2016; 11:e0152317.

Sahnane N, Frattini M, Bernasconi B, Zappa F, Schiavone G, Wannesson L, Antonelli P, Balzarini P, Sessa F, Mazzucchelli L, Tibiletti MG and Martin V. EGFR and KRAS mutations in ALK-positive lung adenocarcinomas: Biological and clinical effect. Clin Lung Cancer. 2016; 17:56-61.

Saintigny P, Peng S, Zhang L, Sen B, Wistuba II, Lippman SM, Girard L, Minna JD, Heymach JV and Johnson FM. Global evaluation of eph receptors and ephrins in lung adenocarcinomas identifies EphA4 as an inhibitor of cell migration and invasion. Mol Cancer Ther. 2012; 11:2021-2032.

Sampsonas F, Ryan D, McPhillips D and Breen DP. Molecular testing and personalized treatment of lung cancer. Curr Mol Pharmacol. 2014; 7:22-32.

Sanger F, Nicklen S and Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977; 74:5463-5467.

Scagliotti G, Brodowicz T, Shepherd FA, Zielinski C, Vansteenkiste J, Manegold C, Simms L, Fossella F, Sugarman K and Belani CP. Treatment-by-histology interaction analyses in three phase III trials show superiority of pemetrexed in nonsquamous non-small cell lung cancer. J Thorac Oncol. 2011; 6:64-70.

Schmidt L, Duh FM, Chen F, Kishida T, Glenn G, Choyke P, Scherer SW, Zhuang Z, Lubensky I, Dean M, Allikmets R, Chidambaram A, Bergerheim UR, Feltis JT, Casadevall C, Zamarron A, Bernues M, Richard S, Lips CJ, Walther MM, Tsui LC, Geil L, Orcutt ML, Stackhouse T, Lipan J, Slife L, Brauch H, Decker J, Niehans G, Hughson MD, Moch H, Storkel S, Lerman

MI, Linehan WM and Zbar B. Germline and somatic mutations in the tyrosine kinase domain of the MET proto-oncogene in papillary renal carcinomas. Nat Genet. 1997; 16:68-73.

Schmitt AM and Chang HY. Long noncoding RNAs in cancer pathways. Cancer Cell. 2016; 29:452-463.

Sekido Y. Molecular pathogenesis of malignant mesothelioma. Carcinogenesis. 2013; 34:1413-1419.

Sequist LV, Gettinger S, Senzer NN, Martins RG, Janne PA, Lilenbaum R, Gray JE, Iafrate AJ, Katayama R, Hafeez N, Sweeney J, Walker JR, Fritz C, Ross RW, Grayzel D, Engelman JA, Borger DR, Paez G and Natale R. Activity of IPI-504, a novel heat-shock protein 90 inhibitor, in patients with molecularly defined non-small-cell lung cancer. J Clin Oncol. 2010; 28:4953-4960.

Sequist LV, Heist RS, Shaw AT, Fidias P, Rosovsky R, Temel JS, Lennes IT, Digumarthy S, Waltman BA, Bast E, Tammireddy S, Morrissey L, Muzikansky A, Goldberg SB, Gainor J, Channick CL, Wain JC, Gaissert H, Donahue DM, Muniappan A, Wright C, Willers H, Mathisen DJ, Choi NC, Baselga J, Lynch TJ, Ellisen LW, Mino-Kenudson M, Lanuti M, Borger DR, Iafrate AJ, Engelman JA and Dias-Santagata D. Implementing multiplexed genotyping of non-small-cell lung cancers into routine clinical practice. Ann Oncol. 2011a; 22:2616-2624.

Sequist LV, Waltman BA, Dias-Santagata D, Digumarthy S, Turke AB, Fidias P, Bergethon K, Shaw AT, Gettinger S, Cosper AK, Akhavanfard S, Heist RS, Temel J, Christensen JG, Wain JC, Lynch TJ, Vernovsky K, Mark EJ, Lanuti M, Iafrate AJ, Mino-Kenudson M and Engelman JA. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. Sci Transl Med. 2011b; 3:75ra26.

Sequist LV, Yang JC, Yamamoto N, O'Byrne K, Hirsh V, Mok T, Geater SL, Orlov S, Tsai CM, Boyer M, Su WC, Bennouna J, Kato T, Gorbunova V, Lee KH, Shah R, Massey D, Zazulina V, Shahidi M and Schuler M. Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR mutations. J Clin Oncol. 2013; 31:3327-3334.

Sharma M, Batra J, Mabalirajan U, Sharma S, Nagarkatti R, Aich J, Sharma SK, Niphadkar PV and Ghosh B. A genetic variation in inositol polyphosphate 4 phosphatase A enhances susceptibility to asthma. Am J Respir Crit Care Med. 2008; 177:712-719.

Sharma SV, Bell DW, Settleman J and Haber DA. Epidermal growth factor receptor mutations in lung cancer. Nat Rev Cancer. 2007; 7:169-181.

Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, McDermott U, Azizian N, Zou L, Fischbach MA, Wong KK, Brandstetter K, Wittner B, Ramaswamy S, Classon M and Settleman J. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell. 2010; 141:69-80.

Shaw AT, Yeap BY, Mino-Kenudson M, Digumarthy SR, Costa DB, Heist RS, Solomon B, Stubbs H, Admane S, McDermott U, Settleman J, Kobayashi S, Mark EJ, Rodig SJ, Chirieac LR, Kwak EL, Lynch TJ and Iafrate AJ. Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK. J Clin Oncol. 2009; 27:4247-4253.

Shaw AT and Solomon B. Targeting anaplastic lymphoma kinase in lung cancer. Clin Cancer Res. 2011; 17:2081-2086.

Shaw AT, Kim DW, Nakagawa K, Seto T, Crino L, Ahn MJ, De Pas T, Besse B, Solomon BJ, Blackhall F, Wu YL, Thomas M, O'Byrne KJ, Moro-Sibilot D, Camidge DR, Mok T, Hirsh V, Riely GJ, Iyer S, Tassell V, Polli A, Wilner KD and Janne PA. Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. N Engl J Med. 2013; 368:2385-2394.

Shaw AT, Engelman JA. ALK in lung cancer: past, present, and future. J Clin Oncol. 2013; 31:1105-1111.

Shaw AT, Kim DW, Mehra R, Tan DS, Felip E, Chow LQ, Camidge DR, Vansteenkiste J, Sharma S, De Pas T, Riely GJ, Solomon BJ, Wolf J, Thomas M, Schuler M, Liu G, Santoro A, Lau YY, Goldwasser M, Boral AL and Engelman JA. Ceritinib in ALK-rearranged non-small-cell lung cancer. N Engl J Med. 2014; 370:1189-1197.

Shaw AT, Gandhi L, Gadgeel S, Riely GJ, Cetnar J, West H, Camidge DR, Socinski MA, Chiappori A, Mekhail T, Chao BH, Borghaei H, Gold KA, Zeaiter A, Bordogna W, Balas B, Puig O, Henschel V, Ou SH and study investigators. Alectinib in ALK-positive, crizotinib-resistant, non-small-cell lung cancer: A single-group, multicentre, phase 2 trial. Lancet Oncol. 2016; 17:234-242.

Shen W, Xi H, Zhang K, Cui J, Li J, Wang N, Wei B and Chen L. Prognostic role of EphA2 in various human carcinomas: A meta-analysis of 23 related studies. Growth Factors. 2014; 32:247-253.

Shi X, Sun M, Liu H, Yao Y and Song Y. Long non-coding RNAs: A new frontier in the study of human diseases. Cancer Lett. 2013; 339:159-166.

Shieh JM, Tang YA, Yang TH, Chen CY, Hsu HS, Tan YH, Salgia R and Wang YC. Lack of association of C-met-N375S sequence variant with lung cancer susceptibility and prognosis. Int J Med Sci. 2013; 10:988-994.

Shigematsu H, Lin L, Takahashi T, Nomura M, Suzuki M, Wistuba II, Fong KM, Lee H, Toyooka S, Shimizu N, Fujisawa T, Feng Z, Roth JA, Herz J, Minna JD and Gazdar AF. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst. 2005; 97:339-346.

Sioris T, Husgafvel-Pursiainen K, Karjalainen A, Anttila S, Kannio A, Salo JA, Perhoniemi V, Heikkila L and Vainio H. Survival in operable non-small-cell lung cancer: Role of p53 mutations, tobacco smoking and asbestos exposure. Int J Cancer. 2000; 86:590-594.

Skov BG, Hogdall E, Clementsen P, Krasnik M, Larsen KR, Sorensen JB, Skov T and Mellemgaard A. The prevalence of EGFR mutations in non-small cell lung cancer in an unselected Caucasian population. APMIS. 2015; 123:108-115.

Smith PM, Elson JL, Greaves LC, Wortmann SB, Rodenburg RJ, Lightowlers RN, Chrzanowska-Lightowlers ZM, Taylor RW and Vila-Sanjurjo A. The role of the mitochondrial ribosome in human disease: Searching for mutations in 12S mitochondrial rRNA with high disruptive potential. Hum Mol Genet. 2014; 23:949-967.

Smits AJ, Kummer JA, Hinrichs JW, Herder GJ, Scheidel-Jacobse KC, Jiwa NM, Ruijter TE, Nooijen PT, Looijen-Salamon MG, Ligtenberg MJ, Thunnissen FB, Heideman DA, de Weger RA and Vink A. EGFR and KRAS mutations in lung carcinomas in the dutch population: Increased EGFR mutation frequency in malignant pleural effusion of lung adenocarcinoma. Cell Oncol (Dordr). 2012; 35:189-196.

Sobin LH, Gospodarowicz MK, Wittekind C. Eds. The TNM Classification of malignant tumours, 7th edition. 2009.

Socinski MA, Goldman J, El-Hariry I, Koczywas M, Vukovic V, Horn L, Paschold E, Salgia R, West H, Sequist LV, Bonomi P, Brahmer J, Chen LC, Sandler A, Belani CP, Webb T, Harper H, Huberman M, Ramalingam S, Wong KK, Teofilovici F, Guo W and Shapiro GI. A multicenter phase II study of ganetespib monotherapy in patients with genotypically defined advanced non-small cell lung cancer. Clin Cancer Res. 2013; 19:3068-3077.

Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, Fujiwara S, Watanabe H, Kurashina K, Hatanaka H, Bando M, Ohno S, Ishikawa Y, Aburatani H, Niki T, Sohara Y, Sugiyama Y and Mano H. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature. 2007; 448:561-566.

Solomon BJ, Mok T, Kim DW, Wu YL, Nakagawa K, Mekhail T, Felip E, Cappuzzo F, Paolini J, Usari T, Iyer S, Reisman A, Wilner KD, Tursi J, Blackhall F and PROFILE 1014 Investigators. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. N Engl J Med. 2014; 371:2167-2177.

Sordella R, Bell DW, Haber DA and Settleman J. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. Science. 2004; 305:1163-1167.

Soria JC, Felip E, Cobo M, Lu S, Syrigos K, Lee KH, Goker E, Georgoulias V, Li W, Isla D, Guclu SZ, Morabito A, Min YJ, Ardizzoni A, Gadgeel SM, Wang B, Chand VK, Goss GD and LUX-Lung 8 Investigators. Afatinib versus erlotinib as second-line treatment of patients with advanced squamous cell carcinoma of the lung (LUX-lung 8): An open-label randomised controlled phase 3 trial. Lancet Oncol. 2015; 16:897-907.

Sos ML, Koker M, Weir BA, Heynck S, Rabinovsky R, Zander T, Seeger JM, Weiss J, Fischer F, Frommolt P, Michel K, Peifer M, Mermel C, Girard L, Peyton M, Gazdar AF, Minna JD, Garraway LA, Kashkar H, Pao W, Meyerson M and Thomas RK. PTEN loss contributes to erlotinib resistance in EGFR-mutant lung cancer by activation of akt and EGFR. Cancer Res. 2009; 69:3256-3261.

Stahl S, Kaminskyy VO, Efazat G, Hyrslova Vaculova A, Rodriguez-Nieto S, Moshfegh A, Lewensohn R, Viktorsson K and Zhivotovsky B. Inhibition of ephrin B3-mediated survival signaling contributes to increased cell death response of non-small cell lung carcinoma cells after combined treatment with ionizing radiation and PKC 412. Cell Death Dis. 2013; 4:e454.

Steels E, Paesmans M, Berghmans T, Branle F, Lemaitre F, Mascaux C, Meert AP, Vallot F, Lafitte JJ and Sculier JP. Role of p53 as a prognostic factor for survival in lung cancer: A systematic review of the literature with a meta-analysis. Eur Respir J. 2001; 18:705-719.

Stinchcombe TE. Novel agents in development for advanced non-small cell lung cancer. Ther Adv Med Oncol. 2014; 6:240-253.

Stone E, Allen HA, Saghaie T, Abbott A, Daniel R, Mead RS, Kohonen-Corish M, Plit M and Morgan L. High proportion of rare and compound epidermal growth factor receptor mutations in an australian population of non-squamous non-small-cell lung cancer. Intern Med J. 2014; 44:1188-1192.

Stransky N, Cerami E, Schalm S, Kim JL and Lengauer C. The landscape of kinase fusions in cancer. Nat Commun. 2014; 5:4846.

Sudo H, Tsuji AB, Sugyo A, Kohda M, Sogawa C, Yoshida C, Harada YN, Hino O and Saga T. Knockdown of COPA, identified by loss-of-function screen, induces apoptosis and suppresses tumor growth in mesothelioma mouse model. Genomics. 2010; 95:210-216.

Sulonen AM, Ellonen P, Almusa H, Lepisto M, Eldfors S, Hannula S, Miettinen T, Tyynismaa H, Salo P, Heckman C, Joensuu H, Raivio T, Suomalainen A and Saarela J. Comparison of solution-based exome capture methods for next generation sequencing. Genome Biol. 2011; 12:R94-2011-12-9-r94.

Sun Y, Yu X, Shi X, Hong W, Zhao J and Shi L. Correlation of survival and EGFR mutation with predominant histologic subtype according to the new lung adenocarcinoma classification in stage IB patients. World J Surg Oncol. 2014; 12:148-7819-12-148.

Suzuki M, Wada H, Yoshino M, Tian L, Shigematsu H, Suzuki H, Alaa M, Tamura H, Fujiwara T, Nagato K, Motohashi S, Moriya Y, Hoshino H, Yoshida S, Shibuya K, Hiroshima K, Nakatani Y and Yoshino I. Molecular characterization of chronic obstructive pulmonary disease-related non-small cell lung cancer through aberrant methylation and alterations of EGFR signaling. Ann Surg Oncol. 2010; 17:878-888.

Takeuchi K, Choi YL, Togashi Y, Soda M, Hatano S, Inamura K, Takada S, Ueno T, Yamashita Y, Satoh Y, Okumura S, Nakagawa K, Ishikawa Y and Mano H. KIF5B-ALK, a novel fusion oncokinase identified by an immunohistochemistry-based diagnostic system for ALK-positive lung cancer. Clin Cancer Res. 2009; 15:3143-3149.

Tan CS, Cho BC and Soo RA. Next-generation epidermal growth factor receptor tyrosine kinase inhibitors in epidermal growth factor receptor -mutant non-small cell lung cancer. Lung Cancer. 2016; 93:59-68.

Tengs T, Lee JC, Paez JG, Zhao X, LaFramboise T, Giannoukos G and Thomas RK. A transforming MET mutation discovered in non-small cell lung cancer using microarray-based resequencing. Cancer Lett. 2006; 239:227-233.

Testa JR, Cheung M, Pei J, Below JE, Tan Y, Sementino E, Cox NJ, Dogan AU, Pass HI, Trusa S, Hesdorffer M, Nasu M, Powers A, Rivera Z, Comertpay S, Tanji M, Gaudino G, Yang H and Carbone M. Germline BAP1 mutations predispose to malignant mesothelioma. Nat Genet. 2011; 43:1022-1025.

Thatcher N, Hirsch FR, Luft AV, Szczesna A, Ciuleanu TE, Dediu M, Ramlau R, Galiulin RK, Balint B, Losonczy G, Kazarnowicz A, Park K, Schumann C, Reck M, Depenbrock H, Nanda S, Kruljac-Letunic A, Kurek R, Paz-Ares L, Socinski MA and SQUIRE Investigators. Necitumumab plus gemcitabine and cisplatin versus gemcitabine and cisplatin alone as first-line therapy in patients with stage IV squamous non-small-cell lung cancer (SQUIRE): An open-label, randomised, controlled phase 3 trial. Lancet Oncol. 2015; 16:763-774.

Thelwell N, Millington S, Solinas A, Booth J and Brown T. Mode of action and application of scorpion primers to mutation detection. Nucleic Acids Res. 2000; 28:3752-3761.

Thorvaldsdottir H, Robinson JT and Mesirov JP. Integrative genomics viewer (IGV): High-performance genomics data visualization and exploration. Brief Bioinform. 2013; 14:178-192.

Thress KS, Paweletz CP, Felip E, Cho BC, Stetson D, Dougherty B, Lai Z, Markovets A, Vivancos A, Kuang Y, Ercan D, Matthews SE, Cantarini M, Barrett JC, Janne PA and Oxnard GR. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. Nat Med. 2015; 21:560-562.

Thunnissen E, van der Oord K and den Bakker M. Prognostic and predictive biomarkers in lung cancer. A review. Virchows Arch. 2014; 464:347-358.

Tiainen M, Tammilehto L, Rautonen J, Tuomi T, Mattson K and Knuutila S. Chromosomal abnormalities and their correlations with asbestos exposure and survival in patients with mesothelioma. Br J Cancer. 1989; 60:618-626.

Tissot C, Couraud S, Tanguy R, Bringuier PP, Girard N and Souquet PJ. Clinical characteristics and outcome of patients with lung cancer harboring BRAF mutations. Lung Cancer. 2016; 91:23-28.

Travis WD, Brambilla E, Müller-Hermelink HK and Harris CC. World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart. Lyon 2004: IARC Press, 2004.

Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger K, Yatabe Y, Powell CA, Beer D, Riely G, Garg K, Austin JH, Rusch VW, Hirsch FR, Jett J, Yang PC, Gould M and American Thoracic Society. International association for the study of lung Cancer/American thoracic Society/European respiratory society: International multidisciplinary classification of lung adenocarcinoma: Executive summary. Proc Am Thorac Soc. 2011; 8:381-385.

Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JH, Beasley MB, Chirieac LR, Dacic S, Duhig E, Flieder DB, Geisinger K, Hirsch FR, Ishikawa Y, Kerr KM, Noguchi M, Pelosi G, Powell CA, Tsao MS, Wistuba I and WHO Panel. The 2015 world health organization classification of lung tumors: Impact of genetic, clinical and radiologic advances since the 2004 classification. J Thorac Oncol. 2015; 10:1243-1260.

Tsao AS, Scagliotti GV, Bunn PA,Jr, Carbone DP, Warren GW, Bai C, de Koning HJ, Yousaf-Khan AU, McWilliams A, Tsao MS, Adusumilli PS, Rami-Porta R, Asamura H, Van Schil PE, Darling GE, Ramalingam SS, Gomez DR, Rosenzweig KE, Zimmermann S, Peters S, Ignatius Ou SH, Reungwetwattana T, Janne PA, Mok TS, Wakelee HA, Pirker R, Mazieres J, Brahmer JR, Zhou Y, Herbst RS, Papadimitrakopoulou VA, Redman MW, Wynes MW, Gandara DR, Kelly RJ, Hirsch FR and Pass HI. Scientific advances in lung cancer 2015. J Thorac Oncol. 2016; 11:613-638.

Tuna M, Machado AS and Calin GA. Genetic and epigenetic alterations of microRNAs and implications for human cancers and other diseases. Genes Chromosomes Cancer. 2016; 55:193-214.

Tuomi T. Fibrous minerals in the lungs of mesothelioma patients: Comparison between data on SEM, TEM, and personal interview information. Am J Ind Med. 1992; 21:155-162.

Tuononen K, Maki-Nevala S, Sarhadi VK, Wirtanen A, Ronty M, Salmenkivi K, Andrews JM, Telaranta-Keerie AI, Hannula S, Lagstrom S, Ellonen P, Knuuttila A and Knuutila S. Comparison of targeted next-generation sequencing (NGS) and real-time PCR in the detection of EGFR, KRAS, and BRAF mutations on formalin-fixed, paraffin-embedded tumor material of non-small cell lung carcinoma-superiority of NGS. Genes Chromosomes Cancer. 2013; 52:503-511.

Tyner JW, Fletcher LB, Wang EQ, Yang WF, Rutenberg-Schoenberg ML, Beadling C, Mori M, Heinrich MC, Deininger MW, Druker BJ and Loriaux MM. MET receptor sequence variants R970C and T992I lack transforming capacity. Cancer Res. 2010; 70:6233-6237.

Ugurluer G, Chang K, Gamez ME, Arnett AL, Jayakrishnan R, Miller RC and Sio TT. Genome-based mutational analysis by next generation sequencing in patients with malignant pleural and peritoneal mesothelioma. Anticancer Res. 2016; 36:2331-2338.

Ulivi P, Chiadini E, Dazzi C, Dubini A, Costantini M, Medri L, Puccetti M, Capelli L, Calistri D, Verlicchi A, Gamboni A, Papi M, Mariotti M, De Luigi N, Scarpi E, Bravaccini S, Turolla GM, Amadori D, Crino L and Delmonte A. Nonsquamous, non-small-cell lung cancer patients who carry a double mutation of EGFR, EML4-ALK or KRAS: Frequency, clinical-pathological characteristics, and response to therapy. Clin Lung Cancer. 2015; [in press].

Uramoto H, Iwata T, Onitsuka T, Shimokawa H, Hanagiri T and Oyama T. Epithelial-mesenchymal transition in EGFR-TKI acquired resistant lung adenocarcinoma. Anticancer Res. 2010; 30:2513-2517.

Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, Zeng K, Malek JA, Costa G, McKernan K, Sidow A, Fire A and Johnson SM. A high-resolution, nucleosome position map

of C. elegans reveals a lack of universal sequence-dictated positioning. Genome Res. 2008; 18:1051-1063.

Van Den Broeck A, Brambilla E, Moro-Sibilot D, Lantuejoul S, Brambilla C, Eymin B, Khochbin S and Gazzeri S. Loss of histone H4K20 trimethylation occurs in preneoplasia and influences prognosis of non-small cell lung cancer. Clin Cancer Res. 2008; 14:7237-7245.

van der Wekken AJ, Saber A, Hiltermann TJ, Kok K, van den Berg A and Groen HJ. Resistance mechanisms after tyrosine kinase inhibitors afatinib and crizotinib in non-small cell lung cancer, a review of the literature. Crit Rev Oncol Hematol. 2016; 100:107-116.

van Dijk EL, Auger H, Jaszczyszyn Y and Thermes C. Ten years of next-generation sequencing technology. Trends Genet. 2014; 30:418-426.

van Meerbeeck JP, Gaafar R, Manegold C, Van Klaveren RJ, Van Marck EA, Vincent M, Legrand C, Bottomley A, Debruyne C, Giaccone G, European Organisation for Research and Treatment of Cancer Lung Cancer Group and National Cancer Institute of Canada. Randomized phase III study of cisplatin with or without raltitrexed in patients with malignant pleural mesothelioma: An intergroup study of the European Organisation for Research and treatment of cancer lung cancer group and the National Cancer Institute of Canada. J Clin Oncol. 2005; 23:6881-6889.

van Zandwijk N, Clarke C, Henderson D, Musk AW, Fong K, Nowak A, Loneragan R, McCaughan B, Boyer M, Feigen M, Currow D, Schofield P, Nick Pavlakis BI, McLean J, Marshall H, Leong S, Keena V and Penman A. Guidelines for the diagnosis and treatment of malignant pleural mesothelioma. J Thorac Dis. 2013; 5:E254-307.

Villa C, Cagle PT, Johnson M, Patel JD, Yeldandi AV, Raj R, DeCamp MM and Raparia K. Correlation of EGFR mutation status with predominant histologic subtype of adenocarcinoma according to the new lung adenocarcinoma classification of the international association for the study of lung Cancer/American thoracic Society/European respiratory society. Arch Pathol Lab Med. 2014; 138:1353-1357.

Vineis P and Wild CP. Global cancer patterns: Causes and prevention. Lancet. 2014; 383:549-557.

Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA,Jr and Kinzler KW. Cancer genome landscapes. Science. 2013; 339:1546-1558.

Vogelzang NJ, Rusthoven JJ, Symanowski J, Denham C, Kaukel E, Ruffie P, Gatzemeier U, Boyer M, Emri S, Manegold C, Niyikiza C and Paoletti P. Phase III study of pemetrexed in combination with cisplatin versus cisplatin alone in patients with malignant pleural mesothelioma. J Clin Oncol. 2003; 21:2636-2644.

Vosa U, Vooder T, Kolde R, Vilo J, Metspalu A and Annilo T. Meta-analysis of microRNA expression in lung cancer. Int J Cancer. 2013; 132:2884-2893.

Wagner JC, Sleggs CA and Marchand P. Diffuse pleural mesothelioma and asbestos exposure in the North Western Cape province. Br J Ind Med. 1960; 17:260-271.

Wang X, Christiani DC, Wiencke JK, Fischbein M, Xu X, Cheng TJ, Mark E, Wain JC and Kelsey KT. Mutations in the p53 gene in lung cancer are associated with cigarette smoking and asbestos exposure. Cancer Epidemiol Biomarkers Prev. 1995; 4:543-548.

Warren GW and Singh AK. Nicotine and lung cancer. J Carcinog. 2013; 12:1-3163.106680.

Watson JD and Crock FH. The structure of DNA. Cold Spring Harb Symp Quant Biol. 1953; 18:123-131.

Wen D, Danquah M, Chaudhary AK and Mahato RI. Small molecules targeting microRNA for cancer therapy: Promises and obstacles. J Control Release. 2015; 219:237-247.

Whitcombe D, Theaker J, Guy SP, Brown T and Little S. Detection of PCR products using self-probing amplicons and fluorescence. Nat Biotechnol. 1999; 17:804-807.

Worzfeld T and Offermanns S. Semaphorins and plexins as therapeutic targets. Nat Rev Drug Discov. 2014; 13:603-621.

Wu JY, Yu CJ, Chang YC, Yang CH, Shih JY and Yang PC. Effectiveness of tyrosine kinase inhibitors on "uncommon" epidermal growth factor receptor mutations of unknown clinical significance in non-small cell lung cancer. Clin Cancer Res. 2011a; 17:3812-3821.

Wu C, Hu Z, He Z, Jia W, Wang F, Zhou Y, Liu Z, Zhan Q, Liu Y, Yu D, Zhai K, Chang J, Qiao Y, Jin G, Liu Z, Shen Y, Guo C, Fu J, Miao X, Tan W, Shen H, Ke Y, Zeng Y, Wu T and Lin D. Genome-wide association study identifies three new susceptibility loci for esophageal squamous-cell carcinoma in chinese populations. Nat Genet. 2011b; 43:679-684.

Wu XC, Zheng YF, Tang M, Li XF, Zeng R and Zhang JR. Association between smoking and p53 mutation in oesophageal squamous cell carcinoma: A meta-analysis. Clin Oncol (R Coll Radiol). 2015; 27:337-344.

Xu CH, Wang W, Wei Y, Hu HD, Zou J, Yan J, Yu LK, Yang RS and Wang Y. Prognostic value of the new international association for the study of lung Cancer/American thoracic Society/European respiratory society classification in stage IB lung adenocarcinoma. Eur J Surg Oncol. 2015; 41:1430-1436.

Yang H, Testa JR and Carbone M. Mesothelioma epidemiology, carcinogenesis, and pathogenesis. Curr Treat Options Oncol. 2008; 9:147-157.

Yang JC, Wu YL, Schuler M, Sebastian M, Popat S, Yamamoto N, Zhou C, Hu CP, O'Byrne K, Feng J, Lu S, Huang Y, Geater SL, Lee KY, Tsai CM, Gorbunova V, Hirsh V, Bennouna J, Orlov S, Mok T, Boyer M, Su WC, Lee KH, Kato T, Massey D, Shahidi M, Zazulina V and Sequist LV. Afatinib versus cisplatin-based chemotherapy for EGFR mutation-positive lung adenocarcinoma (LUX-lung 3 and LUX-lung 6): Analysis of overall survival data from two randomised, phase 3 trials. Lancet Oncol. 2015; 16:141-151.

Yasuda H, Park E, Yun CH, Sng NJ, Lucena-Araujo AR, Yeo WL, Huberman MS, Cohen DW, Nakayama S, Ishioka K, Yamaguchi N, Hanna M, Oxnard GR, Lathan CS, Moran T, Sequist LV, Chaft JE, Riely GJ, Arcila ME, Soo RA, Meyerson M, Eck MJ, Kobayashi SS and Costa DB. Structural, biochemical, and clinical characterization of epidermal growth factor receptor (EGFR) exon 20 insertion mutations in lung cancer. Sci Transl Med. 2013; 5:216ra177.

Ye K, Schulz MH, Long Q, Apweiler R and Ning Z. Pindel: A pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics. 2009; 25:2865-2871.

Yeddula N, Xia Y, Ke E, Beumer J and Verma IM. Screening for tumor suppressors: Loss of ephrin receptor A2 cooperates with oncogenic KRas in promoting lung adenocarcinoma. Proc Natl Acad Sci U S A. 2015; 112:E6476-85.

Yoon JH, Smith LE, Feng Z, Tang M, Lee CS and Pfeifer GP. Methylated CpG dinucleotides are the preferential targets for G-to-T transversion mutations induced by benzo[a]pyrene diol epoxide in mammalian cells: Similarities with the p53 mutation spectrum in smoking-associated lung cancers. Cancer Res. 2001; 61:7110-7117.

Yoon S, Woo SU, Kang JH, Kim K, Shin HJ, Gwak HS, Park S and Chwae YJ. NF-kappaB and STAT3 cooperatively induce IL6 in starved cancer cells. Oncogene. 2012; 31:3467-3481.

Yoshikawa Y, Sato A, Tsujimura T, Emi M, Morinaga T, Fukuoka K, Yamada S, Murakami A, Kondo N, Matsumoto S, Okumura Y, Tanaka F, Hasegawa S, Nakano T and Hashimoto-Tamaoki T.

Frequent inactivation of the BAP1 gene in epithelioid-type malignant mesothelioma. Cancer Sci. 2012; 103:868-874.

Yu J, Bulk E, Ji P, Hascher A, Tang M, Metzger R, Marra A, Serve H, Berdel WE, Wiewroth R, Koschmieder S and Muller-Tidow C. The EPHB6 receptor tyrosine kinase is a metastasis suppressor that is frequently silenced by promoter DNA hypermethylation in non-small cell lung cancer. Clin Cancer Res. 2010; 16:2275-2283.

Yu HA, Arcila ME, Rekhtman N, Sima CS, Zakowski MF, Pao W, Kris MG, Miller VA, Ladanyi M and Riely GJ. Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. Clin Cancer Res. 2013; 19:2240-2247.

Yun CH, Mengwasser KE, Toms AV, Woo MS, Greulich H, Wong KK, Meyerson M and Eck MJ. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. Proc Natl Acad Sci U S A. 2008; 105:2070-2075.

Zalcman G, Mazieres J, Margery J, Greillier L, Audigier-Valette C, Moro-Sibilot D, Molinier O, Corre R, Monnet I, Gounant V, Riviere F, Janicot H, Gervais R, Locher C, Milleron B, Tran Q, Lebitasy MP, Morin F, Creveuil C, Parienti JJ, Scherpereel A and French Cooperative Thoracic Intergroup (IFCT). Bevacizumab for newly diagnosed pleural mesothelioma in the mesothelioma avastin cisplatin pemetrexed study (MAPS): A randomised, controlled, open-label, phase 3 trial. Lancet. 2016; 387:1405-1414.

Zauderer MG, Bott M, McMillan R, Sima CS, Rusch V, Krug LM and Ladanyi M. Clinical characteristics of patients with malignant pleural mesothelioma harboring somatic BAP1 mutations. J Thorac Oncol. 2013; 8:1430-1433.

Zhang Z, Wang T, Zhang J, Cai X, Pan C, Long Y, Chen J, Zhou C and Yin X. Prognostic value of epidermal growth factor receptor mutations in resected non-small cell lung cancer: A systematic review with meta-analysis. PLoS One. 2014; 9:e106053.

Zhao Y, Yang J, Chen Z, Gao Z, Zhou F, Li X, Li J, Shi S, Feng X, Sun N, Yao R, Zhou C, Chang S, Li M, Zhou Y, Li L, Zhang X and He J. Identification of somatic alterations in stage I lung adenocarcinomas by next-generation sequencing. Genes Chromosomes Cancer. 2014; 53:289-298.

Zhao F, Xu M, Lei H, Zhou Z, Wang L, Li P, Zhao J and Hu P. Clinicopathological characteristics of patients with non-small-cell lung cancer who harbor EML4-ALK fusion gene: A meta-analysis. PLoS One. 2015; 10:e0117333.

Zhou C, Wu YL, Chen G, Feng J, Liu XQ, Wang C, Zhang S, Wang J, Zhou S, Ren S, Lu S, Zhang L, Hu C, Hu C, Luo Y, Chen L, Ye M, Huang J, Zhi X, Zhang Y, Xiu Q, Ma J, Zhang L and You C. Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): A multicentre, open-label, randomised, phase 3 study. Lancet Oncol. 2011; 12:735-742.

Zhuang G, Song W, Amato K, Hwang Y, Lee K, Boothby M, Ye F, Guo Y, Shyr Y, Lin L, Carbone DP, Brantley-Sieders DM and Chen J. Effects of cancer-associated EPHA3 mutations on lung cancer. J Natl Cancer Inst. 2012; 104:1182-1197.

ORIGINAL PUBLICATIONS

Recent Publications in this Series

39/2016 Alexey Yukin
Animal Models of Early Brain Disorders: Behavioural and Electrophysiological Approaches
40/2016 Satu Valo
Western Diet and Genetic Predisposition as Risk Factors of Colon Cancer
41/2016 Riccardo Siligato
Hormonal Regulation of Primary and Secondary Growth in the Root of *Arabidopsis thaliana*
42/2016 Janne Tynell
Virus-Host Interactions of Emerging Respiratory Pathogens
43/2016 Katri Korpela
Intestinal Microbiota Development in Childhood: Implications for Health and Disease
44/2016 Mikko Muona
Identification of New Genetic Syndromes with Epilepsy by Whole-Exome Sequencing
45/2016 Ana Cathia Dias Magalhães
On the Mechanisms of Neural Development in the Ventral Telencephalon
46/2016 Petri Arvilommi
Treatment, Adherence and Disability in Bipolar Disorder
47/2016 Annika Thomson
Finnish Pretrial Male Firesetters: Mortality, Suicidality, Psychopathy and Morbidity of Schizophrenia
48/2016 Sami Svanbäck
Toward Accurate High-Throughput Physicochemical Profiling using Image-Based Single-Particle Analysis
49/2016 Katarzyna Leskinen
The Roles of YbeY, RfaH, and Hfq in Gene Regulation and Virulence of *Yersinia enterocolitica* O:3
50/2016 Arjan van Adrichem
Discovery of Small Molecule Modulators of Ras Superfamily Proteins Studies of MgcRacGAP and Ras
51/2016 Meri Kekäle
Chronic Myeloid Leukemia Patients' Adherence to Tyrosine Kinase Inhibitors in Finland – A Journey of Eighty-six Patients
52/2016 Jaakko Kopra
Endogenous GDNF as a Regulator of Midbrain Dopamine Neurons
53/2016 Vasili Stegajev
Histamine Receptor 4 (H4R) In pathogenesis of Sjögren's Syndrome
54/2016 Katriina Viitasalo
Shift Work and Cardiovascular Risk Factors — Prevention among Airline Employees
55/2016 Michal Stepniewski
Computational Studies on Membrane Proteins and Membrane-Drug Interactions
56/2016 Jaana Oikkonen
Genetics and Genomics of Musical Abilities
57/2016 Hanna Antila
Activation of the TrkB Neurotrophin Receptor by Antidepressant Drugs
58/2016 Elina Rydman
Inflammatory Effects of Nanosized Titanium Dioxide and Carbon Nanotube Pulmonary Exposure

Helsinki 2016 ISSN 2342-3161 ISBN 978-951-51-2426-5